Optimal Cross Hedging of Insurance Derivatives using Quadratic BSDEs

by

Ludovic Tangpi Ndounkeu

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Mathematics at Stellenbosch University

Department of Mathematical Sciences
University of Stellenbosch
Private Bag X1, 7602 Matieland, South Africa

Supervisor: Dr. Raouf Ghomrasni

December 2011

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....
L. Tangpi

Date: 2011/05/02

Abstract

We consider the utility portfolio optimization problem of an investor whose activities are influenced by an exogenous financial risk (like bad weather or energy shortage) in an incomplete financial market. We work with a fairly general non-Markovian model, allowing stochastic correlations between the underlying assets. This important problem in finance and insurance is tackled by means of backward stochastic differential equations (BSDEs), which have been shown to be powerful tools in stochastic control. To lay stress on the importance and the omnipresence of BSDEs in stochastic control, we present three methods to transform the control problem into a BSDEs. Namely, the martingale optimality principle introduced by Davis, the martingale representation and a method based on Itô-Ventzell's formula. These approaches enable us to work with portfolio constraints described by closed, not necessarily convex sets and to get around the classical duality theory of convex analysis. The solution of the optimization problem can then be simply read from the solution of the BSDE. An interesting feature of each of the different approaches is that the generator of the BSDE characterizing the control problem has a quadratic growth and depends on the form of the set of constraints. We review some recent advances on the theory of quadratic BSDEs and its applications. There is no general existence result for multidimensional quadratic BSDEs. In the one-dimensional case, existence and uniqueness strongly depend on the form of the terminal condition. Other topics of investigation are measure solutions of BSDEs, notably measure solutions of BSDE with jumps and numerical approximations. We extend the equivalence result of Ankirchner *et al.* (2009) between existence of classical solutions and existence of measure solutions to the case of BSDEs driven by a Poisson process with a bounded terminal condition. We obtain a numerical scheme to approximate measure solutions. In fact, the existing self-contained construction of measure solutions gives rise to a numerical scheme for some classes of Lipschitz BSDEs. Two numerical schemes for quadratic BSDEs introduced in Imkeller *et al.* (2010) and based, respectively, on the Cole-Hopf transformation and the truncation procedure are implemented and the results are compared.

Keywords: BSDE, quadratic growth, measure solutions, martingale theory, numerical scheme, indifference pricing and hedging, non-tradable underlying, defaultable claim, utility maximization.

Opsomming

Ons beskou die nuts portefeulje optimalisering probleem van 'n belegger wat se aktiwiteite beïnvloed word deur 'n eksterne finansiële risiko (soos onweer of 'n energie tekort) in 'n onvolledige finansiële mark. Ons werk met 'n redelik algemene nie-Markoviaanse model, wat stogastiese korrelasies tussen die onderliggende bates toelaat. Hierdie belangrike probleem in finansies en versekering is aangepak deur middel van terugwaartse stogastiese differensiaalvergelykings (TSDEs), wat blyk om 'n onderskeidende metode in stogastiese beheer te wees. Om klem te lê op die belangrikheid en alomteenwoordigheid van TSDEs in stogastiese beheer, bespreek ons drie metodes om die beheer probleem te transformeer na 'n TSDE. Naamlik, die martingale optimaliteits beginsel van Davis, die martingale voorstelling en 'n metode wat gebaseer is op 'n formule van Itô-Ventzell. Hierdie benaderings stel ons in staat om te werk met portefeulje beperkinge wat beskryf word deur geslote, nie noodwendig konvekse versamelings, en die klassieke dualiteit teorie van konvekse analise te oorkom. Die oplossing van die optimaliserings probleem kan dan bloot afgelees word van die oplossing van die TSDE. 'n Interessante kenmerk van elkeen van die verskillende benaderings is dat die voortbringer van die TSDE wat die beheer probleem beshryf, kwadratiese groei en afhanglik is van die vorm van die versameling beperkings. Ons herlei 'n paar onlangse vooruitgange in die teorie van kwadratiese TSDEs en gepaartgaande toepassings. Daar is geen algemene bestaanstelling vir multidimensionele kwadratiese TSDEs nie. In die een-dimensionele geval is bestaan ââen uniekheid sterk afhanklik van die vorm van die terminale voorwaardes. Ander ondersoek onderwerpe is maatoplossings van TSDEs, veral maatoplossings van TSDEs met spronge en numeriese benaderings. Ons brei uit op die ekwivalensie resultate van Ankirchner *et al.* (2009) tussen die bestaan van klassieke oplossings en die bestaan van maatoplossings vir die geval van TSDEs wat gedryf word deur 'n Poisson proses met begrensde terminale voorwaardes. Ons verkry 'n numeriese skema om oplossings te benader. Trouens, die bestaande self-vervatte konstruksie van maatoplossings gee aanleiding tot 'n numeriese skema vir sekere klasse van Lipschitz TSDEs. Twee numeriese skemas vir kwadratiese TSDEs, bekendgestel in Imkeller *et al.* (2010), en gebaseer is, onderskeidelik, op die Cole-Hopf transformasie en die afknot proses is geïmplementeer en die resultate word vergelyk.

Acknowledgements

The Almighty is my Lord and my guide, I thank him for everything. I wholeheartedly thank my supervisor Raouf Ghomrasni for his dedication and all his support. His sagacious advises and his availability have greatly improved this thesis. Raouf gave me a topic that I have enjoyed working on, and has helped me to acquire skills that will exalt my work for the years to come. I am grateful to Frances Aron, Rhoda Hawkins and Douw Steyn who read my draft and introduced me to the art of scientific writing. Thank you to Lafras Uys for the translation of the abstract. It is unfortunate that my Afrikaans is still very poor.

I express my heartfelt appreciation to the African Institute for Mathematical Sciences for the financial and multifaceted support and I thank each one of its staff members. AIMS provided me with a nice place for work and enabled me to engage with some of the finest students and lecturers that the world has to offer. The discussions I had with the people who visited the AIMS Research Centre, and with my colleagues have helped me growing and learning; I am thankful to them. Further, for the peaceful and serene atmosphere around me, for helping me to go through this, I thank my friends, especially Jeanne and my officemates Mihaja and Tahiri. Finally, I would like to express my deepest love to my parents and sisters. Thank you for always being there for me, thank you for your love and your care, thank you for trusting me so much.

Dedications

To my mom ...

Contents

Chapter 1

Brownian Model of Cross Hedging

Named after the physicist Robert Brown, Brownian motion is the mathematical concept used to describe the motion of particles suspended in a fluid due to thermally driven molecular collision. The work of Robert Merton and Paul Samuelson laid the foundations of what is now known as Brownian motion models of financial markets. In these models, financial instruments such as assets, gains or portfolios are modelled by stochastic processes driven by Brownian motion. It is an extension of the one-period and discrete-time model of Markowitz.

1.1 Quadratic Hedging: Basic Concepts

In this section, we introduce some basic concepts and terminologies that will be used throughout this chapter.

1.1.1 Residual Risk, Basis Risk

In general, insurance and financial products are for materials or tangible underlyings, and a trader owning a contract written on a given asset will invest in a portfolio containing shares of the asset to cover himself against a loss linked to the contract. What happens if the trader is exposed to a risk based on a non-tradable underlying? For instance, an investor owning an industry of umbrellas will expect a wet winter to sell lots of umbrellas, but is exposed to the risk of having a dry winter instead. Thus, a clever attitude should be to buy an insurance contract which pays a certain amount of money if it does not rain a lot (a weather derivative for instance). Yet, it is not possible to buy some shares of rainfall. The question for both the buyer and the seller of such a contract is how to price and to hedge it. A good way to deal with this could be to invest in a tradable asset which is strongly correlated to a sort of rainfall index, this could be for instance a production of corn. The new asset is called the hedging instrument, and the first one (the rainfall index) is the hedged asset. The investor, by using a hedging instrument, exposes himself to a residual risk.

Definition 1.1.1. *A residual risk is any risk remaining to an investment after all other risks have been eliminated, or hedged.*

In other words, when hedging a non-tradable asset, if the hedging instrument is imperfectly correlated with the asset that carries the risk, then there is a part of the risk which is not hedged. This is the residual risk. An example of derivatives which could evoke a residual risk is the index option. This is a financial derivative tied to the price of stock market indices, and it is generally hedged by trading only some of the underlyings, and this leads to a residual risk. More generally, in practice, hedging risk cannot be eliminated totally by hedging with futures contracts. This could be because the future is often not perfectly correlated to the risk the investor bears, or because the hedged asset is different from the hedging instrument. We define the *basis* as the difference between the price of the hedged asset and the price of the hedged instrument. In Ankirchner and Imkeller (2011), the residual risk is also referred to as *basis risk*. When hedging financial derivatives, we should use a technique which minimizes the residual risk.

1.1.2 Characterization of the Hedging

Before introducing some quantities characterising the (mean-variance) hedging based on Ankirchner and Imkeller (2011), let us briefly recall the concept of correlation.

Let X and Y be two random variables. The degree of correlation between X and Y is measured by the correlation coefficient defined by the formula

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

The coefficient ρ lies between -1 and 1. If $\rho = \pm 1$ then we have a "perfect correlation". If $\rho = 0$ then X and Y are not related at all (we say they are independent), and the closer ρ is to 1 (or -1), the more X and Y are related.

Usually in mathematical finance, we deal with stochastic processes. When we talk about correlation of two stochastic processes, we mean the instantaneous correlation process $(\rho_t)_{t\in[0,T]}$ defined by $\rho_t = \frac{Cov(X_t,Y_t)}{\sqrt{Var(X_t)}\sqrt{Var(Y_t)}}$, for a given time t in the finite trading time interval $t \in [0,T]$. In most of the financial models dealing with correlation, the instantaneous correlation coefficient is assumed to be constant. Note that it can also be assumed to be a (deterministic) function of t, or for more realistic models, a stochastic process itself.

Let us now assume that a trader wants to hedge an asset Y with the hedging instrument X. We assume further that there exists a future contract written on X with price P at time 0 and maturity T, which is also the delivery time of Y. Let Y_T and X_T, be the prices at time T of the assets Y and X, respectively. The trader has invested in N_Y assets Y, and he buys k future contracts on N_X assets X. At time T the value of one future contract is $N_X X_T$, thus at the horizon the trader earns (or looses) $N_X X_T - P$ from a single contract. Consequently, the total spending of the investor is $N_Y Y_T - k(N_X X_T - P)$, with variance

$$V = E\left[\left(N_Y Y_T - k(N_X X_T - P)\right)^2\right] - (E\left[N_Y Y_T - k(N_X X_T - P)\right])^2.$$

After some reductions, and by using the formula $\rho = \frac{E[Y_T X_T] - E[Y_T]E[X_T]}{\sigma_X \sigma_Y}$, where σ_X^2 and σ_Y^2 are the variance of Y_T and X_T, respectively, we have

$$V = K(N_Y^2 \sigma_Y^2 - 2kN_Y N_X \rho \sigma_X \sigma_Y + k^2 N_X^2 \sigma_X^2).$$

Where K is a constant depending on the sample size. The variance is minimum if the number k of futures is

$$k^* = \rho \frac{N_Y}{N_X} \frac{\sigma_Y}{\sigma_X}.$$

This leads to the following definition.

Definition 1.1.2. *The* hedge ratio *is the number k^* of futures that the trader needs to buy in order to minimize the variance.*

According to Ankirchner and Imkeller (2011), the factor N_X adjusts the units of the futures to the quantity of assets Y needed, and the factor $\rho\frac{\sigma_Y}{\sigma_X}$ determines the proportion of risk on Y that should be transferred to X in order to minimize the variance.

In order to define the hedging error, let us consider the simple case of static hedging, with $\Delta X = X_T - X_0$, and $\Delta Y = Y_T - Y_0$ two standard Gaussian variables that are strongly correlated, with correlation coefficient ρ.

To protect himself from having to pay $N_Y \Delta Y$, the trader holds k futures, which will produce the expected amount $kN_X \Delta X$. The quadratic hedging consists in minimizing the quadratic error

$$E\left[(N_Y \Delta Y - kN_X \Delta X)^2\right].$$

There exists a normally distributed random variable N, independent to ΔX such that

$$\Delta Y = \sqrt{1-\rho^2} N + \rho \Delta X.$$

Multiplying both sides by ΔX, and using the fact that $E[\Delta X^2] = 1$ and $E[\Delta X \Delta Y] = \rho$, we have that $E[\Delta X N] = 0$. Furthermore, the quadratic error becomes

$$E\left[\left(\rho N_Y \Delta X + N_Y \sqrt{1-\rho^2} N - kN_X \Delta X\right)^2\right] = (\rho N_Y - kN_X)^2 + N_Y(1-\rho^2),$$

which is minimum if, and only if,

$$k = \rho \frac{N_Y}{N_X}.$$

For this value of k, the quadratic error is $N_Y \sqrt{1-\rho^2} N$. This quantity is called *hedging error.*

Thus, the hedging error is a non-increasing function of the basis given by $\sqrt{1-\rho^2}$. The quantity $\sqrt{1-\rho^2}$ represents the percentage contribution of the basis to the total error. This means for instance that if the correlation is $\rho = 0.945$, the basis would represent 32.7% of the hedging error, thus only 62,3% of the price of Y can be hedged. The following graph shows the relationship between the correlation coefficient and the percentage of the price of Y that can be hedged. The graph shows that the percentage of the price that can be hedged increases as the correlation increases. However, (as also pointed out by Ankirchner and Imkeller (2011)) the observation of the above graph displays some drawbacks of the mean variance hedging of a non-tradable asset with a correlated one. When the correlation is high, a small variation of the correlation leads to a very big variation of the hedged percentage, whereas for small correlations, a small variation of the correlation will lead to almost no change of the percentage. Furthermore, even for high correlations (but of course strictly less than 1), the hedged percentage is still not significantly high.

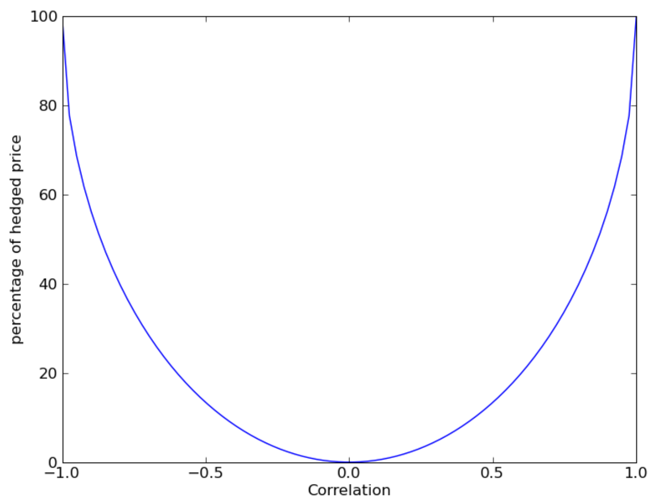


Figure 1.1: Percentage of hedged price for varying correlation.

1.2 Utility Based Hedging

The problem of hedging of financial derivatives in a complete financial market setting under the no arbitrage hypothesis has a definitive answer. For example, by Delbaen and Schachermayer (1994), it is obtained by the martingale representation, under the unique equivalent martingale measure, of the price process. However in an incomplete market, it is not usually possible to perfectly hedge derivatives. Besides the quadratic hedging presented in Section 1.1, many approaches attempting to solve this problem have been provided. Among which we have the superreplication of El Karoui and Quenez (1995), which gives an interval $[p_{buy}, p_{sell}]$ where the "fair" price of the derivative should lie to have an arbitrage free market, where p_{buy} is the superreplicating price of the buyer and p_{sell} the superreplicating price of the seller. We also have the method mixing option hedging and utility maximization introduced by Hodges and Neuberger (1989), who first used the theory of utility to problems of decision in mathematical finance. In Section 1.1, we described some drawbacks of the quadratic hedging of derivatives written on non-tradable underlyings. The method was static, i.e. after the investment has been done at time 0, no investments is done in between the initial time and the horizon time, and the method was rather simple because the trader dealt with random variables (X_T, Y_T) but not with the whole processes of prices $((X_t)_{t\in[0,T]}, (Y_t)_{t\in[0,T]})$. We shall present in the rest of the thesis a dynamic and more sophisticated method of hedging, which allows the trader to invest continuously according to a hedging strategy. The method will require to solve a stochastic control problem.

1.2.1 Setting of the Model

We consider a Brownian motion model for financial market with probability space $(\Omega, \mathcal{F}, P)$ carrying a 2-dimensional standard Brownian motion $W = (W_t)_{t\in[0,T]}$, where T is a fixed strictly positive real number and $[0, T]$ the time interval. The flow of information is given by the filtration $\mathbb{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}$, which is the augmented filtration defined in the following way.

Definition 1.2.1. *Let $\{\mathcal{F}_t^W, 0 \leq t \leq T\}$ the natural filtration generated by $(W_t)_{t\in[0,T]}$ with for all t, $\mathcal{F}_t^W = \sigma(\{W_s, 0 \leq s \leq t\})$. The augmented filtration (or augmented Brownian filtration) $\mathbb{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}$ is defined by $\mathcal{F}_t = \sigma(\mathcal{F}_t^W \cup \mathcal{N})$, where $\mathcal{N} = \{E \subset \Omega; \exists G \in \mathcal{F}, E \subset G, P(G) = 0\}$ is the set of P-negligible sets.*

It is shown, see Karatzas and Shreve (1988), Corollary 2.7.8, that the filtration $\mathbb{F}$ is continuous, and that, see Theorem 2.7.9 of the same reference, W is still a Brownian motion with respect to $\mathbb{F}$. Unless otherwise stated, measurable functions will be measurable with respect to the sigma algebra $\mathcal{F}$, by adapted or predictable processes we shall refer to $\mathbb{F}$-adapted (or $\mathbb{F}$-predictable), and by almost surely (a.s.) we mean with respect to the probability P.

Let Y be a (non-tradable) financial instrument, which we shall call the hedged asset, based on an external risk and modelled by the Itô-diffusion

$$dY_t = a(t, Y_t)\, dt + b(t, Y_t)\, dW_t^1, \quad \text{and} \quad Y_0 = y_0.$$

With the assumption that $a : [0, T] \times \mathbb{R} \to \mathbb{R}$ and $b : [0, T] \times \mathbb{R} \to \mathbb{R}$ are two Borel measurable deterministic functions satisfying the Lipschitz and linear growth conditions, i.e. there exists a constant $C \in \mathbb{R}_+$ such that for all $t \in [0, T]$ and $x, x' \in \mathbb{R}$

$$|a(t,x) - a(t,x')| + |b(t,x) - b(t,x')| \leq C|x - x'| \tag{1.2.1}$$

$$|a(t,x)| + |b(t,x)| \leq C(1 + |x|). \tag{1.2.2}$$

Let F be a derivative written on Y, where F is a bounded[1] $\mathcal{F}_T$-random variable. We further assume to have a financial market which allows investments and short positions on a risk-less bank account and a risky asset. To lighten the notation, the risk-less asset is used as numéraire, which is equivalent to taking its instantaneous interest rate as 0 at every time. Define the process

$$dB_t := \rho_t \, dW_t^1 + \sqrt{1 - \rho_t^2} \, dW_t^2. \tag{1.2.3}$$

By the Lévy characterization of Brownian motion, $(B_t)_{t \in [0,T]}$ is a Brownian motion. Moreover, B and W^1 are correlated with instantaneous correlation coefficient $\rho_t \in [-1, 1]$. The price process of the risky asset is modelled by the dynamics

$$dS_t = S_t(\mu_t \, dt + \sigma_t \, dB_t),$$

where $(\mu_t)_{t \in [0,T]}$ and $(\sigma_t)_{t \in [0,T]}$ are two predictable processes with $\sigma_t > 0$ for all t. The price process and the external risk Y are correlated (through the Brownian motion). The correlation is a deterministic function of time t. The drift μ and the volatility σ are $\mathbb{F}$-predictable processes assumed to be uniformly bounded.

Some authors who study the issue of correlated financial assets have other ways to model the correlation between the assets. On the one hand, Ankirchner *et al.* (2008), model the correlation with two independent Brownian motions W and B, but in their settings B drives the dynamics of the risk process, while the price of the risky asset that the trader invests in is driven by both Brownian motions and the risk process affects the drift part of the price process. In the other hand, the model of Ankirchner *et al.* (2010*b*) has only one Brownian motion, and the volatility matrices are assumed to be correlated.

The trader is aiming to hedge the risk that the possession of the derivative F involves. In that regard, he invests in the above described financial market. He is a small investor, and his action will not influence the movement of the prices in the market. Derivatives such as F, written on a non-tradable underlying, are called *insurance derivatives*; and this technique of hedging an insurance derivative via a correlated asset is known as *cross hedging*. The first insurance derivatives were offered by Chicago Mercantile Exchange in 1997, it was a future contract on accumulated heating degree days (cHDD). If κ is the temperature above which rooms are heated, and τ the average temperature of the day, the cHDD is the average sum of heating degree days (HDDs

[1] The study of Chapter 3 will show that this boundedness assumption is crucial in our model.

$= \max\{0, \kappa - \tau\}$). This kind of financial products are generaly bought by energy and agricultural compagnies because they can easily be affected by dry winters or wet summers. Here is a well known example of insurance derivative that our model can be applied to.

Example 1.2.2 (Weather derivative)**.** *There are many types of weather derivatives, this is an instance which is pretty much like an European option. A financial agent with weather exposure can choose to invest in a future (swap) contract which makes him earn money if the degree days within the "trading period" are greater than a fixed threshold, and which makes him pay the counterpart if the degree days do not exceed the threshold.*

The trader decides at time $t \in [0, T]$ what amount π_t of (current) wealth to invest in the risky asset. The number of shares is thus given by the formula $\frac{\pi_t}{S_t}$. The one-dimensional process $\pi = (\pi_t)_{t\in[0,T]}$ is called the investment strategy, and is the process which controls the overall wealth of the investor. It is a nonanticipative process, which means that π is progressively measurable. We assume that the investor chooses his investment strategies based on some constraints. He might for instance impose a threshold that his total wealth must not exceed, or decides not to take any loans. We are therefore in the situation of a constrained investment problem. We summarize the constraints by assuming that a strategy should be in a given set C to fulfil all the required constraints of the investor. Unlike most of the works dealing with constrained investment problems, we do not assume C to be convex. Instead, we follow the path of Hu *et al.* (2005) and assume the set C to be closed. Albeit it is a less restrictive assumption, it will provide a key argument, see Remark 2.3.5.

Define $x > 0$ the initial wealth of the investor. His wealth X_t^π at time t, if he runs the strategy $(\pi_t)_{t\in[0,T]}$ is given by

$$\begin{aligned} X_t^\pi &= x + \int_0^t \frac{\pi_s}{S_s}\, dS_s \\ &= x + \int_0^t \pi_s \sigma_s(\theta_s\, ds + dB_s); \end{aligned} \tag{1.2.4}$$

where $\theta = \frac{\mu}{\sigma}$ is a uniformly bounded process called the market price of risk.

Let U be the utility function of the investor. In this thesis, unless otherwise stated, the utility function is a deterministic function $U : \mathbb{R}_+ \to \mathbb{R}$ which is continuously differentiable, strictly increasing and strictly concave, satisfying the Inada conditions:

$$U'(0) = \lim_{x\to 0} U'(x) = \infty \quad \text{and} \quad U'(\infty) = \lim_{x\to\infty} U'(x) = 0,$$

and with asymptotic elasticity strictly less than 1, i.e.

$$AE(U) = \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} < 1.$$

Definition 1.2.3. *An admissible trading strategy is a one-dimensional process* $\pi = (\pi_t)_{t\in[0,T]}$ *such that*

- π *is progressively measurable*
- $E\left[\int_0^t |\pi_s\sigma_s|^2\, ds\right] < \infty$ *a.s., for all* $t \in [0,T]$
- π *is self-financing*
- $\pi \in C\ \lambda \otimes P$ *a.s.*[2]
- $\{U(X^\pi_\tau) : \tau\, \text{stopping time}\}$ *is uniformly integrable.*

The set of admissible trading strategies is denoted $\mathcal{A}$*.*

Moreover, we shall consider in some cases that at any time $t \in [0,T]$ the investor consumes the non-negative stochastic wealth c_t. The (consumption) process $(c_t)_{t\in[0,T]}$ forms with the investment process $(\pi_t)_{t\in[0,T]}$ the strategy of the investor. The admissibility conditions on the consumption process are that $(c_t)_{t\in[0,T]}$ should be predictable, integrable, and should belong $\lambda \otimes P$ a.s. to a constraint set C' also assumed to be closed. We shall use the same notation $\mathcal{A}$ to denote the set of admissible strategies (π, c). The wealth process takes the form

$$X_t^{\pi,c} = x + \int_0^t \pi_s\sigma_s(\theta_s\, ds + dB_s) - \int_0^t c_s\, ds \tag{1.2.5}$$

$$= x + \int_0^t X_s^{\pi,c}\tilde{\pi}_s\sigma_s(\theta_s\, ds + dB_s) - \int_0^t X_s^{\pi,c}\tilde{c}_s\, ds, \tag{1.2.6}$$

where $\tilde{\pi}$ and $\tilde{c}$ are, respectively the fraction of money invested in the risky asset and the fraction of money consumed, i.e. $\tilde{\pi} = \pi/X^{\pi,c}$ and $\tilde{c} = c/X^{\pi,c}$. The set of admissible strategies does not contain arbitrage opportunities. In fact, consider the probability[3] measure defined by $Q = \exp\left(-\int_0^T \theta_s^2\, ds - \int_0^T \theta_s\, dB_s\right)\cdot P$, we have $Q \sim P$. Let (π, c) be an admissible strategy. Under Q, the process $X^{\pi,c}$ is a supermartingale, due to Girsanov's theorem, Doob-Meyer decomposition of supermartingales and the fact that c_t is non-negative for all t. Thus, if $X_0^{\pi,c} = 0$ then $E^Q\left[X_T^{\pi,c}|\mathcal{F}_0\right] \leq X_0^{\pi,c}$ implies $E^Q\left[X_T^{\pi,c}\right] = 0$. Since Q is equivalent to P, we conclude that the set of admissible strategies is free of arbitrage. We will often use the following notation. For a given function $g(\tilde{c})$ and a vector a,

$$dist(a, C) = ess\inf_{\pi\in C} |a - \pi|, \quad \max_{\tilde{c}\in C'} g(\tilde{c}) = ess\sup_{\tilde{c}\in C'} g(\tilde{c}).$$

[2] λ is the Lebesgue measure defined on the σ-algebra of Borel sets $\mathcal{B}_{[0,T]}$.

[3] Since θ is uniformly bounded, it follows from Novikov's criterion that Q is indeed a probability.

The maximal expected utility of the investor is

$$V^F(x) = \sup_{\pi \in \mathcal{A}} E\left[\int_0^T \alpha U(c_t)\,dt + U(X_T^\pi - F)|\mathcal{F}_0\right];$$

when he starts at $t = 0$ with $X_0 = x$, and pays out the liability F at the horizon. The parameter α is a positive constant defined by the investor, and we will always put $\alpha = 0$ if the consumption in not taken into account. One of the concern of our study is to solve the following dynamical version of the expected utility maximization problem:

$$V^F(t, x_t) = \sup_{\pi \in \mathcal{A}} E\left[\int_t^T \alpha U(c_s)\,ds + U(X_T^{x_t,\pi} - F)|\mathcal{F}_t\right] \tag{1.2.7}$$

where

$$X_T^{x_t,\pi} = x_t + \int_t^T \frac{\pi_s}{S_s} dS_s - \int_t^T c_s\,ds$$

is the terminal wealth if the investor starts at time $t \in [0, T]$ with the wealth x_t. We call V^F defined by (1.2.7) the *value function*, and (1.2.7) is known as the utility indifference hedging problem. We should add, in addition, that the financial market is incomplete, because the risk cannot be perfectly hedged, and due to constraints in the choice of the strategies in this model with finite horizon not every contract F is perfectly attainable.

This control problem —used here to cross hedge a derivative written on an illiquid or non-tradable underlying— is the usual formulation of the problem of indifference pricing which aims at finding the value $h_t(x_t)$ of the future at time t that makes the investor indifferent between trading with initial wealth x_t at time t and paying nothing at the horizon, and trading with initial wealth $x_t + h_t(x_t)$ at time t and paying F at the horizon. The indifference value is thus implicitly defined by

$$V^0(t, x_t) = V^F(t, x_t + h_t(x_t)).$$

This class of problem has given rise to a wealth of literature, and authors have presented three main approaches that the large majority of papers related to the problem try to improve by relaxing some assumptions or taking into account other factors to make the model more realistic. The Ph.D. thesis of Frei (2009) extensively comments on these three approaches in its introduction.

The first group of papers, in a Markovian setting, solve the problem by the HJB equation. This will be briefly done in Section 2.1. The second group of papers deal with duality theory. The usual method in this case is, under Brownian filtration framework, to transform the dynamic problem (1.2.7) into a static one, and then use the optional decomposition theorem of supermartingales to express the constraints of the static problem in terms of the density of the equivalent martingale measures (which will be infinitely many). This allows the construction of the dual analogue of the static problem. In cases of constrained problems like ours, the constraint set C needs in general to be convex, what we do not assume here.

The third group of papers give a characterisation of the control by a BSDE. The first step in this method is usually to apply the martingale optimality principle to show that the value function can be obtained via the solution of a BSDE. As explained in Frei (2009), this step is somehow similar to the HJB theory, though more general because it does not require the Markovian assumption. Then, the issue of existence and uniqueness of solution of the BSDE needs to be addressed. El Karoui *et al.* (1997) explicitly precise the link between BSDEs and finance. The BSDE approach for problems such as Problem (1.2.7) is of growing interest, especially since the beginning of the last decade. We mention among works dealing with some of its aspects, Hu *et al.* (2005), where the problem is treated for different utility functions in a multi-dimensional setting and with closed set of constraints. In a similar but more recent work, Cheridito and Hu (2010) characterize the optimal consumption and investment when the consumption process must also lie in a closed set of constraints, but without stochastic correlation. Becherer (2006) proves a BSDE characterization of the control process π. More recently, Frei introduced in his Ph.D. thesis the assumption of stochastic correlation. Ankirchner *et al.* (2010a), are concerned with the case where F is a defaultable contingent claim and the market model allows a random jump; and Frei and Reis (2011) discuss the existence of a “Nash equilibrium", when dealing with the relative performance of interacting traders, in the sense of having simultaneous optimal strategies for all traders. In most of the cases, the BSDE derived by mean of the martingale optimality principle is of quadratic growth, i.e. its generator grows quadratically (see Definition 3.1.1). This has stimulated research in the field because of the need to provide a general result of existence and uniqueness for this class of equations. The next chapter studies martingale optimality, or the transition from the optimization problem to the BSDE.

Chapter 2

Martingale Optimality Principle in Control

Since the works of M. Davis who introduced the martingale optimality principle in the 1970s, martingale methods have been used in stochastic control. Hu *et al.* (2005) stimulated the use of the martingale optimality principle in stochastic finance by showing that the principle can help to transform a control problem into a BSDE. Hence, it provides a fully probabilistic technique, alternative to the stochastic maximum principle, to solve a stochastic control problem. In addition this technique enables one to handle non-convex sets of constraints.

2.1 Digression into the Markovian Case

In this section, we make a short digression into the Markovian case to expose how the HJB equation arises and the method to find an analytic solution, both for the cross hedging problem and the optimal investment problem.

Here in this section, we assume μ, σ and ρ to be constant, the claim F is an explicit function of Y_T say $F = \phi(Y_T)$ and the control π is a Markov control, i.e. $\pi_t = g(t, X_t)$ for a given measurable function g from $[0, T] \times \mathbb{R}$ to A, subset of $\mathbb{R}$. Moreover, we take $c = \alpha = 0$, i.e. there is no consumption.

The stochastic control problem (1.2.7) in this setting is given by the formula[1]

$$V(t, x, y) = \sup_{\pi \in \mathcal{A}} E\big[U(X_T - \phi(Y_T)) \;\big|\; X_t = x, Y_t = y\big]. \tag{2.1.1}$$

[1]We shall use the notation $E[Z|X_t = x, Y_t = y] = E_{t,x,y}[Z]$.

Assuming $V \in C^2([0,T]\times\mathbb{R}^2,\mathbb{R})$, Itô's formula yields

$$\begin{aligned}V(\theta,X_\theta,Y_\theta)-V(t,X_t,Y_t) = &\int_t^\theta \big(V_t+\pi_s\mu V_x + a(s,Y_s)V_y \\ &+\frac{1}{2}(\pi_s^2\sigma^2 V_{xx}+b^2(s,Y_s)V_{yy})+\rho\pi_s\sigma b(s,Y_s)V_{xy}\big)\,ds \\ &+\int_t^\theta \pi_s\sigma V_x\,dB_s+\int_t^\theta b(s,Y_s)V_y\,dW_s.\end{aligned}$$

Let us recall two results.

Theorem 2.1.1 (Dynamic Programming Principle). *The value function of the control Problem* (2.1.1) *solves*

$$V(t,x,y)=\sup_{\pi\in\mathcal{A}} E_{t,x,y}\left[V(\theta,X_\theta,Y_\theta)\right], \tag{2.1.2}$$

with $(t,x,y)\in[0,T]\times\mathbb{R}^2$ and $\theta\in[t,T]$.

Proof. See Yong and Zhou (1999) Theorem 3.3 page 180. □

Proposition 2.1.2. *The infinitesimal generator of the two-dimensional Itô-diffusion* (X,Y) *defined by*

$$\begin{cases} dX_t=\pi_t\mu\,dt+\pi_t\sigma\,dB_t\\ dY_t=a(t,Y_t)\,dt+b(t,Y_t)\,dW_t^1\end{cases}$$

is given by $\mathcal{L}^\pi$*, with*

$$\mathcal{L}^\pi f=\frac{\partial f}{\partial t}+\pi_t\mu\frac{\partial f}{\partial x}+a(t,y)\frac{\partial f}{\partial y}+\frac{1}{2}\left(\pi_t^2\sigma^2\frac{\partial^2 f}{\partial^2 x}+b^2(t,y)\frac{\partial^2 f}{\partial^2 y}\right)+\rho\pi_t\sigma b(t,y)\frac{\partial f}{\partial x\partial y}.$$

Proof. See Tangpi (2010). □

From (2.1.2), we have

$$\begin{aligned}V(t,x,y)=\sup_{\pi\in\mathcal{A}}E_{t,x,y}\Big[&V(t,X_t,Y_t)+\int_t^\theta \big(V_t+\pi_s\mu V_x+a(s,Y_s)V_y\\ &+\frac{1}{2}(\pi_s^2\sigma^2V_{xx}+b^2(s,Y_s)V_{yy})+\rho\pi_s\sigma b(s,Y_s)V_{xy}\big)\,ds\\ &+\int_t^\theta \pi_s\sigma V_x\,dW_s+\int_t^\theta b(s,Y_s)V_y\,dB_s\Big].\end{aligned}$$

Taking out the martingale part, we obtain

$$\sup_{\pi\in\mathcal{A}}E_{t,x,y}\left[\int_t^\theta \mathcal{L}^\pi V(s,X_s,Y_s)\,ds\right]=0.$$

Using the fact that π is a Markov control, multiplying by $\frac{1}{\theta - t}$, and applying a limit argument, we are led to the HJB equation

$$V_t + \sup_{\pi \in A} \left\{ \pi_t \mu V_x + \frac{1}{2}\pi_t^2 \sigma^2 V_{xx} + \rho \pi_t \sigma b(t,y) V_{xy} \right\} + a(t,y)V_y + \frac{1}{2} b^2(t,y) V_{yy} = 0, \tag{2.1.3}$$

with terminal condition $V(T,x,y) = U(x - \phi(y))$.

Let us assume that the trader has a power utility $U(x) = \frac{x^\gamma}{\gamma}$, $\gamma \in (0,1)$.

Theorem 2.1.3. *The trader's value function is given by*

$$V(t,x,y) = \frac{x^\gamma}{\gamma} E_{t,y} \left[\frac{(X_T - \phi(Y_T))^{\gamma(1-\rho^2)}}{x^{\gamma(1-\rho^2)}} \exp\left\{ \frac{1}{2}(1-\rho^2)\frac{\mu^2}{\sigma^2}\frac{\gamma}{\gamma - 1}(T-t) \right\} \right]^{\frac{1}{1-\rho^2}}.$$

Proof. We make the *ansatz* $V(t,x,y) = \frac{x^\gamma}{\gamma} h(t,y)$. Substituting in (2.1.3) yields

$$\begin{aligned} x^2 h_t(t,y) + \gamma \sup_{\pi \in A} \left\{ \frac{1}{2}\pi_t^2 \sigma^2 (\gamma - 1) h(t,y) + \pi_t (\mu x h(t,y) + \rho \sigma b(t,y) x h_y(t,y)) \right\} \\ + x^2 a(t,y) h_y(t,y) + \frac{1}{2} b^2(t,y) x^2 h_{yy} = 0. \end{aligned} \tag{2.1.4}$$

We find the optimizer

$$\pi_t^* = -\frac{\mu x h(t,y) + \rho \sigma b(t,y) x h_y(t,y)}{\sigma^2 (\gamma - 1) h(t,y)}.$$

Plugging it in (2.1.4), the equation becomes

$$h_t - \frac{1}{2}\gamma \frac{(\mu h(t,y) + \rho \sigma b(t,y) h(t,y))^2}{\sigma^2(\gamma - 1)h(t,y)} + a(t,y) h_y(t,y) + \frac{1}{2} b^2(t,y) h_{yy} = 0,$$

with terminal condition $h(T,y) = \frac{(x-\phi(y))^\gamma}{x^\gamma}$. Now we use the *distortion method* to transform this last non-linear PDE into a linear one. The distortion method, introduced by Zariphopoulou (2001), consists of making the power transformation

$$h(t,y) = u(t,y)^{\frac{1}{1-\rho^2}}. \tag{2.1.5}$$

This leads to

$$u_t + \frac{1}{2} b^2(t,y) u_{yy} + \left(a(t,y) - \rho \frac{\mu}{\sigma} \frac{\gamma}{\gamma - 1} b(t,y) \right) u_y - \frac{1}{2}(1-\rho^2)\frac{\mu^2}{\sigma^2}\frac{\gamma}{\gamma-1} u = 0,$$

with terminal condition

$$u(T,y) = \frac{(x - \phi(y))^{\gamma(1-\rho^2)}}{x^{\gamma(1-\rho^2)}}.$$

Using Feynman-Kac's representation of solutions of Cauchy problems (see Karatzas and Shreve (1988) Theorem 5.7.6) we have

$$u(t,y) = E_{t,y}\left[\frac{(X_T - \phi(Y_T))^{\gamma(1-\rho^2)}}{x^{\gamma(1-\rho^2)}} \exp\left\{\frac{1}{2}(1-\rho^2)\frac{\mu^2}{\sigma^2}\frac{\gamma}{\gamma-1}(T-t)\right\}\right].$$

Combining the different transformations, we find the formula giving the value function. □

By the verification theorem, the optimal (cross hedging) strategy is given by:

$$\pi_t^* = -\frac{\mu x h(t,y) + \rho\sigma b(t,y) x h_y(t,y)}{\sigma^2(\gamma-1)h(t,y)}.$$

2.1.0.1 Optimal Investment Problem

Still using the HJB theory, the optimal investment problem given by

$$V(t,x,y) = \sup_{\pi\in\mathcal{A}} E_{t,x}\left[U(X_T)\right],$$

that is, with $\phi = 0$ can be solved in a slightly different way, without using Feynman-Kac's formula, but by direct integration of an ordinary differential equation. See for instance Tangpi (2010), where the optimal investment problem is solved for different utility functions. In the case of power utility the value function corresponds to

$$V(t,x) = \frac{x^\gamma}{\gamma}\exp\left(\gamma\left[r + \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2(1-\gamma)}\right]\cdot(T-t)\right).$$

2.2 Martingale Optimality Principle

The martingale optimality principle provides a way to confirm one's guess of the candidate for the control which maximizes the cost functional of a control problem. The idea is to find a functional that is a surpermartingale for every control, but a martingale for the optimal controls.

The martingale optimality principle, due to Davis, is a well known principle in stochastic analysis and more precisely in stochastic control theory. The idea of the principle comes from the definition of a martingale itself: an integrable stochastic process $(X_t)_{t\in[0,T]}$ such that for all $s,t \in [0,T]$ with $t \leq s$, $E\left[X_s|\mathcal{F}_t\right] = X_t$. In other words, the future state of the process (i.e. X_s) is likely to be the same as the current state (i.e. X_t) given the accumulated knowledge we have. Given this, could an investor hope his wealth process to be a martingale? Of course yes, provided that he thinks he is investing optimally. A formal mathematical answer to the question will be given in the following results.

For a pair $(t, x) \in [0, T] \times \mathbb{R}$ taken arbitrary and fixed, consider the stochastic control problem $v(t, x) = \sup_{\pi \in \mathcal{A}} E\left[U(X_T^{\pi,x})|\mathcal{F}_t\right]$ with $X_t^{\pi,x} = x$.

Theorem 2.2.1 (The martingale optimality principle). *If there exists a control strategy π^* such that the function $g(t, x) := E\left[U(X_T^{\pi^*,x})|\mathcal{F}_t\right]$ satisfies:*

1. $(g(s, X_s^{\pi^*,x}))_{s\in[t,T]}$ *is a martingale*
2. $(g(s, X_s^{\pi,x}))_{s\in[t,T]}$ *is a supermartingale for all $\pi \in \mathcal{A}$.*

Then we obtain:

a . π^* is an optimal control strategy

b . For all initial states (t, x) of the controlled process we have $g(t, x) = v(t, x)$, i.e. g coincides with the value function.

Proof. Let $\pi \in \mathcal{A}$, and π^* a control strategy satisfying the hypothesis of the theorem. Notice that for each x fixed in $\mathbb{R}$, $g(T, x) = U(x)$. We have

$$\begin{aligned}
E\left[U(X_T^{\pi^*,x})|\mathcal{F}_t\right] &= E\left[g(T, X_T^{\pi^*,x})|\mathcal{F}_t\right] \\
&= g(t, x) \qquad (2.2.1) \\
&\geq E\left[g(T, X_T^{\pi,x})|\mathcal{F}_t\right] \qquad (2.2.2) \\
&= E\left[U(X_T^{\pi,x})|\mathcal{F}_t\right].
\end{aligned}$$

Thus, π^* is optimal. Equation (2.2.1) comes from the fact that $(g(s, X_s^{\pi^*,x}))_{s\in[t,T]}$ is a martingale and Equation (2.2.2) follows from the supermartingale property of $(g(s, X_s^{\pi,x}))_{s\in[t,T]}$. Moreover, by definition of g we have $g \leq v$ and from the above calculations $g(t, x) \geq v(t, x)$ for each t, x. Hence, g is value function. □

Remark 2.2.2.

- *The reader could find a description of the principle in Korn (2003) where the study is done in the Markovian case, and thus the conditional expectations are not on a σ-algebra, but rather on a fixed state (t, x) of the controlled process.*
- *The martingale optimality principle does not give any suggestion on how to construct an optimal strategy. Even less, it is not an existence results. However, it gives an important criterion of optimality which will help us to get around the classical HJB theory to solve our stochastic control problem in a purely probabilistic way.*
- *One should notice that the function g defined in the previous theorem is a random function, since the conditional expectation is a random variable.*

The principle described above is used to solve a variety of control problems. Rogers and Williams (1987) give a considerable list of control problems and explain how to solve them by means of the martingale optimality. Among papers applying the principle to solve problems related to our's, we can mention the work of Korn and Menkens (2005) that solves the problem of worst-case portfolio optimization (which consists in finding the portfolio with the worst-case expected utility bound when the stock price is subject to uncertain downward jumps). They apply the martingale optimality to derive the HJB equation from the Bellman's principle, see Theorem 2 of the above mentioned reference. We can also quote Yang and Zhang, where in a Markovian model similar to the one described in Section 2.1, but with external risk process allowing random jumps, they prove a verification theorem using the martingale optimality principle, see Yang and Zhang (2005), Theorem 1.

Now we will exploit the fact that the controlled process in our setting is the wealth process, given by (1.2.5) to derive other results. We rewrite the control problem as

$$v(t,x) = \sup_{(\pi,c)\in\mathcal{A}} E\left[U\left(x + \int_t^T \pi_s \frac{dS_s}{S_s} - \int_t^T c_s\, ds\right)|\mathcal{F}_t\right].$$

For the sake of notational simplicity we omit the dependence of X on x.

Corollary 2.2.3. *Let (π^*, c^*) in $\mathcal{A}$ and consider the function defined for each (t,x) taken in $[0,T]\times\mathbb{R}$ by $g(t,x) := E\left[U\left(x + \int_t^T \pi_s^* \frac{dS_s}{S_s} - \int_t^T c_s^*\, ds\right)|\mathcal{F}_t\right]$. Let $(\pi, c) \in \mathcal{A}$, put $Z_s^{\pi,c} = g(s, X_s^{\pi,c})$, $t \le s \le T$. If for all $(\pi,c)\in\mathcal{A}$ $(Z_s^{\pi,c})_{s\in[t,T]}$ is a supermartingale, then (π^*, c^*) is optimal.*

Proof. For all $s \ge t$,

$$\begin{aligned}
&E\left[Z_s^{\pi^*,c^*}|\mathcal{F}_t\right] \\
&\quad = E\left[g(s, x + \int_t^s \pi_u^* \frac{dS_u}{S_u} - \int_t^s c_u^*\, du)|\mathcal{F}_t\right] \\
&\quad = E\left[E\left[U\left(x + \int_t^s \pi_u^* \frac{dS_u}{S_u} - \int_t^s c_u^*\, du + \int_s^T \pi_u^* \frac{dS_u}{S_u} - \int_s^T c_u^*\, du\right)|\mathcal{F}_s\right]|\mathcal{F}_t\right] \\
&\quad = E\left[U\left(x + \int_t^T \pi_u^* \frac{dS_u}{S_u} - \int_t^T c_u^*\, du\right)|\mathcal{F}_t\right] \\
&\quad = g(t,x) = g(t, X_t^{\pi^*,c^*}) = Z_t^{\pi^*,c^*}.
\end{aligned}$$

Hence, $(Z_s^{\pi^*,c^*})_{s\in[t,T]}$ is a martingale. We conclude with Theorem 2.2.1 that (π^*, c^*) is optimal. In addition, the function $(t,x) \mapsto g(t,x)$ is the value function. □

For this particular case where the controlled process is the wealth process, the proof of the above corollary implies that, for the definition of g given above,

$Z^{\pi^*,c*}$ is always a martingale. Note that Corollary 2.2.3 is a stronger result that Theorem 2.2.1.

Define for all $s \in [t,T]$ and $(\pi,c) \in \mathcal{A}$, $Y_s^{\pi,c} = v(s, X_s^{\pi,c})$ and $\mathcal{A}_s(\pi,c)$ the set of admissible strategies that coincide with (π,c) on $[t,s]$, i.e. $(\widehat{\pi},\widehat{c})$ belongs to $\mathcal{A}_s(\pi,c)$ means that $(\widehat{\pi}_u,\widehat{c}_u) = (\pi_u,c_u)$ for all $t \le u \le s$. Note that if $s \le \theta$ then $\mathcal{A}_\theta(\pi,c) \subset \mathcal{A}_s(\widehat{\pi},\widehat{c})$. We are now equipped to state the following optimality principle taken from Mania and Tevzadze (2008).

Proposition 2.2.4. *Let $t \in [0,T]$, $x \in \mathbb{R}$. Assume that $v(t,x) < \infty$, then*

1. *For all $(\pi,c) \in \mathcal{A}$ $(Y_s^{\pi,c})_{s\in[t,T]}$ is a supermartingale*
2. *(π^*,c^*) is optimal if, and only if, $(Y_s^{\pi^*,c^*})_{s\in[t,T]}$ is a martingale.*

Proof. We start by proving the first claim of the proposition. Let $(\widehat{\pi},\widehat{c}) \in \mathcal{A}$, $t \le s \le \theta \le T$.

$$
\begin{aligned}
&E\left[Y_\theta^{\widehat{\pi},\widehat{c}} \mid \mathcal{F}_s\right] \\
&\quad = E\left[v\left(\theta, x + \int_t^\theta \widehat{\pi}_u \frac{dS_u}{S_u} - \int_t^\theta \widehat{c}_u\,du\right) \mid \mathcal{F}_s\right] && (2.2.3)\\
&\quad = E\left[\sup_{(\pi,c)\in\mathcal{A}} E\left[U\left(x + \int_t^\theta \widehat{\pi}_u \frac{dS_u}{S_u} - \int_t^\theta \widehat{c}_u\,du\right.\right.\right. \\
&\qquad\quad \left.\left.\left. + \int_\theta^T \pi_u \frac{dS_u}{S_u} - \int_\theta^T c_u\,du\right) \mid \mathcal{F}_\theta\right] \mid \mathcal{F}_s\right] && (2.2.4)\\
&\quad = E\left[\sup_{(\pi,c)\in\mathcal{A}_\theta(\widehat{\pi},\widehat{c})} E\left[U\left(x + \int_t^T \pi_u \frac{dS_u}{S_u} - \int_t^T c_u\,du\right) \mid \mathcal{F}_\theta\right] \mid \mathcal{F}_s\right] \\
&\quad = \sup_{(\pi,c)\in\mathcal{A}_\theta(\widehat{\pi},\widehat{c})} E\left[U\left(x + \int_t^T \pi_u \frac{dS_u}{S_u} - \int_t^T c_u\,du\right) \mid \mathcal{F}_s\right] && (2.2.5)\\
&\quad \le \sup_{(\pi,c)\in\mathcal{A}_s(\widehat{\pi},\widehat{c})} E\left[U\left(x + \int_t^T \pi_u \frac{dS_u}{S_u} - \int_t^T c_u\,du\right) \mid \mathcal{F}_s\right] && (2.2.6)\\
&\quad = \sup_{(\pi,c)\in\mathcal{A}} E\left[U\left(x + \int_t^s \widehat{\pi}_u \frac{dS_u}{S_u} - \int_t^s \widehat{c}_u\,du + \int_s^T \pi_u \frac{dS_u}{S_u} + \int_s^T c_u du\right) \mid \mathcal{F}_s\right] \\
&\quad = v(s, X_s^{\widehat{\pi},\widehat{c}}) = Y_s^{\widehat{\pi},\widehat{c}}.
\end{aligned}
$$

Therefore, $Y^{\widehat{\pi},\widehat{c}}$ is a supermartingale. To obtain Equation (2.2.5), we use the tower property and the well known fact that if the utility function has an asymptotic elasticity strictly less than 1 then an optimal control exists. Inequality (2.2.6) comes from $\mathcal{A}_\theta(\widehat{\pi},\widehat{c}) \subset \mathcal{A}_s(\widehat{\pi},\widehat{c})$.

Now let us prove the second claim of Proposition 2.2.4. Assume that (π^*, c^*) is optimal. For all $t \le s \le \theta \le T$,

$$\begin{aligned}
&E\left[Y_\theta^{\pi^*,c^*} \mid \mathcal{F}_s\right] \\
&= E\left[v\left(\theta, x + \int_t^\theta \pi^* \frac{dS_u}{S_u} - \int_t^\theta c_u^*\, du\right) \mid \mathcal{F}_s\right] && (2.2.7)\\
&= E\left[\sup_{(\pi,c)\in\mathcal{A}} E\left[U\left(x + \int_t^\theta \pi_u^* \frac{dS_u}{S_u} - \int_t^\theta c_u^* du\right.\right.\right.\\
&\qquad \left.\left.\left. + \int_\theta^T \pi_u \frac{dS_u}{S_u} - \int_\theta^T c_u du\right) \mid \mathcal{F}_\theta\right] \mid \mathcal{F}_s\right] && (2.2.8)\\
&\ge E\left[E\left[U\left(x + \int_t^\theta \pi_u^* \frac{dS_u}{S_u} - \int_t^\theta c_u^*\, du + \int_\theta^T \pi_u^* \frac{dS_u}{S_u} - \int_\theta^T c_u^*\, du\right) \mid \mathcal{F}_\theta\right] \mid \mathcal{F}_s\right]\\
&= E\left[U\left(x + \int_t^T \pi_u^* \frac{dS_u}{S_u} - \int_t^T c_u^*\, du\right) \mid \mathcal{F}_s\right]\\
&\ge \sup_{(\pi,c)\in\mathcal{A}_s(\pi^*,c^*)} E\left[U\left(x + \int_t^T \pi_u \frac{dS_u}{S_u} - \int_t^T c_u\, du\right) \mid \mathcal{F}_s\right]\\
&= \sup_{(\pi,c)\in\mathcal{A}} E\left[U\left(x + \int_t^s \pi_u^* \frac{dS_u}{S_u} - \int_t^s c_u^*\, du + \int_s^T \pi_u^* \frac{dS_u}{S_u} - \int_s^T c_u^*\, du\right) \mid \mathcal{F}_s\right]\\
&= v(s, X_s^{\pi^*,c^*}) = Y_s^{\pi^*,c^*}.
\end{aligned}$$

This means that Y^{π^*,c^*} is a submartingale. We conclude that it is a martingale since it is in addition a supermartingale. The converse is a consequence of both the first claim and Theorem 2.2.1. The reader could find an alternative proof of this proposition in the appendix of Mania and Tevzadze (2008), where the authors do not consider the consumption process, see Mania and Tevzadze (2008) Proposition A.1. □

The previous result gives a criterion of optimality in terms of the value function of the control problem, not in terms of the objective function like Theorem 2.2.1. The next result is a generalization of the Bellman's principle of optimality to a stochastic and non-Markovian system.

Theorem 2.2.5. *Let $t \in [0, T]$, $x \in \mathbb{R}$. Assume that $v(t, x) < \infty$, then for all $s \in [t, T]$*

$$v(t, x) = \sup_{(\pi,c)\in\mathcal{A}} E\left[v\left(s, x + \int_t^s \pi_u \frac{dS_u}{S_u} - \int_t^s c_u\, du\right) \mid \mathcal{F}_t\right]. \qquad (2.2.9)$$

Proof. Let $s \in [t, T]$ and $(\widehat{\pi}, \widehat{c}) \in \mathcal{A}$.

$$E\left[v\left(s, x + \int_t^s \widehat{\pi}_u \frac{dS_u}{S_u} - \int_t^s \widehat{c}_u\, du\right) \mid \mathcal{F}_t\right] \le v(t, x)$$

because $Y^{\widehat{\pi},\widehat{c}}$ is a supermartingale. Hence, taking the supremum, we have

$$\sup_{(\widehat{\pi},\widehat{c})\in\mathcal{A}} E\left[v\left(s, x+\int_t^s \widehat{\pi}_u \frac{dS_u}{S_u} - \int_t^s \widehat{c}_u\, du\right) \mid \mathcal{F}_t\right] \le v(t,x). \tag{2.2.10}$$

Moreover,

$$\begin{aligned} E&\left[v\left(s, x+\int_t^s \widehat{\pi}_u \frac{dS_u}{S_u} - \int_t^s \widehat{c}_u\, du\right) \mid \mathcal{F}_t\right] \\ &\ge E\left[E\left[U\left(x+\int_t^T \widehat{\pi}_u \frac{dS_u}{S_u} - \int_t^T \widehat{c}_u\, du\right) \mid \mathcal{F}_s\right] \mid \mathcal{F}_t\right] \\ &= E\left[U\left(x+\int_t^T \widehat{\pi}_u \frac{dS_u}{S_u} - \int_t^T \widehat{c}_u\, du\right) \mid \mathcal{F}_t\right]. \end{aligned}$$

This follows from the definition of the value function v, the supremum and the tower property of the conditional expectation. Hence, taking the supremum we have

$$v(t,x) \le \sup_{(\pi,c)\in\mathcal{A}} E\left[v\left(s, x+\int_t^s \pi_u \frac{dS_u}{S_u} - \int_t^s c_u\, du\right)\right].$$

This inequality combined with (2.2.10) lead to the claimed result (2.2.9). □

2.3 BSDE Characterizations

In our non-Markovian framework, it can be hard to describe the value function and the optimal strategy by means of the HJB equation. We present in this section some alternative approaches, more general and purely probabilistic. The methods presented use the optimality principles of the previous section, and more generally the martingale theory, to express the value function and the optimal strategies in terms of the solution of a BSDE.

Let us recall a couple of definitions. The concepts of martingales of bounded mean oscillation (BMO-martingales for short) and stochastic exponentials will play a key role in our analysis. The reader may refer to Appendix A for some results from the theory of BMO-martingales.

Definition 2.3.1. *A continuous local martingale of the form $M^Z = \int_0^\cdot Z_s\, dB_s$ is said to be a BMO-martingale if, and only if,*

$$\|M^Z\|_{BMO} = \sup_{\tau,\,\mathbb{F}-stopping\ time} E\left[\int_\tau^T |Z_s|^2\, ds | \mathcal{F}_\tau\right]^{\frac{1}{2}} < \infty.$$

Definition 2.3.2. *The stochastic exponential, also known as Doléans-Dade exponential, of a semimartingale X such that $X_0 = 0$ is the solution of the stochastic integral equation*

$$Y_t = 1 + \int_0^t Y_{s-}\, dX_s.$$

It is denoted by $\mathcal{E}(X)$.

In the particular case where X is a continuous local martingale, it is known from Doléans-Dade, see for example Kazamaki (1994), that $\mathcal{E}(X)$ can be defined as

$$\mathcal{E}(X)_t = \exp\left(X_t - \frac{1}{2}[X]_t\right),$$

where $[X]_t$ is the quadratic variation of X.

Example 2.3.3. *If X is an Itô process of the form $X_t = \int_0^t F_s\, dB_s + \int_0^t G_s\, ds$, then $[X]_t = \int_0^t F_s^2\, ds$ and for $G \equiv 0$, $\mathcal{E}(X) = \exp\left(\int_0^t F_s\, dB_s - \frac{1}{2}\int_0^t F_s^2\, ds\right)$.*

In the next subsection we shall attempt to solve of the control problem considering different utility functions. We start by exponential utility without considering consumption, then we deal with the CRRA utility functions.

2.3.1 Characterization via Martingale Optimality

The method was used by Hu *et al.* (2005), to characterise a stochastic control problem by a BSDE, and they used it based on the observation that the expected exponential utility can be computed using the martingale optimality principle.

Our goal in applying the principle is to construct a family of stochastic processes $K = \left\{K^{\pi,c} = (K_t^{\pi,c})_{t\in[0,T]}\right\}$ endowed with the following properties:

P1. $K_T^{\pi,c} = \int_0^T \alpha U(c_t)\, dt + U\left(X_T^{\pi,c} - F\right)$, for all $(\pi, c) \in \mathcal{A}$

P2. $K_0^{\pi,c} = K_0$ is constant for all $(\pi, c) \in \mathcal{A}$

P3. $K^{\pi,c}$ is a supermartingale for all $(\pi, c) \in \mathcal{A}$

P4. There exists (at least one) $(\pi^*, c^*) \in \mathcal{A}$ for which K^{π^*,c^*} is a martingale.

Constructing such a family K of processes will indeed help to describe the value function and the optimal strategy. On the one hand, the properties **P2** and **P4** of K imply

$$K_0 = E\left[K_T^{\pi^*,c^*}\right] = E\left[U\left(X_T^{\pi^*,c^*} - F\right) + \int_0^T \alpha U(c_t^*)\, dt\right]. \tag{2.3.1}$$

On the other hand, Property **P3** of K implies

$$E\left[U\left(X_T^{\pi,c} - F\right) + \int_0^T \alpha U(c_t)\, dt\right] = E\left[K_T^{\pi,c}\right] \leq E\left[K_0\right].$$

Therefore, $E\left[U\left(X_T^{\pi,c} - F\right) + \int_0^T \alpha U(c_t)\, dt\right] \leq E\left[U\left(X_T^{\pi^*,c^*} - F\right) + \int_0^T \alpha U(c_t^*)\, dt\right]$ and,

$$V^F(x) = K_0, \tag{2.3.2}$$

hence (π^*, c^*) is an optimal strategy (notice that V^F does depend on x, through $X_T^{\pi^*,c^*}$).

At this stage, it would be quite difficult to go any further in the construction of K, because the construction of the processes $K^{\pi,c}$, $(\pi, c) \in \mathcal{A}$ depends on the utility function. Therefore, we shall continue the study in the next subsections for particular cases of utility functions.

Since it reduces the calculations, the exponential utility is a commonly used choice of utility function in cross hedging and indifference valuation. This choice of utility function also has an interesting financial consequence in indifference pricing. It yields an indifference price which does not depend on the initial wealth x.

2.3.1.1 Case of Exponential Utility

We assume that the trader has an exponential utility, given by $U(x) = -e^{-\eta x}$ with $\eta \in (0,1)$, and we take $\alpha = 0$. The results of this subsection are mostly due to Hu *et al.* (2005) and Ankirchner and Imkeller (2011). We add some more detailed proofs.

For all $t \in [0,T]$ and $\pi \in \mathcal{A}$ let

$$K_t^\pi = -\exp\left(-\eta\left(X_t^\pi - Y_t\right)\right);$$

where (Y, Z) is a solution of the BSDE

$$Y_t = F - \int_t^T Z_s\, dB_s - \int_t^T f(s, Z_s)\, ds, \quad t \in [0,T]. \tag{2.3.3}$$

Assume (Y, Z) exists. Let us define the family $M = \left\{M^\pi = (M_t^\pi)_{t\in[0,T]} : \pi \in \mathcal{A}\right\}$ of local martingales by

$$\begin{aligned} M_t^\pi &= \exp\left(-\eta(x - Y_0)\right) \exp\left(-\int_0^t \eta(\pi_s\sigma_s - Z_s)\, dB_s - \frac{1}{2}\int_0^t \eta^2(\pi_s\sigma_s - Z_s)^2\, ds\right) \\ &= \exp\left(-\eta(x - Y_0)\right) \mathcal{E}\left(-\int_0^\cdot \eta(\pi_s\sigma_s - Z_s)\, dB_s\right)_t. \end{aligned} \tag{2.3.4}$$

The following result holds.

Theorem 2.3.4. *Assume that the parameters f and F are such that Equation (2.3.3) has a solution, and that $f(t,z)$ satisfies the condition*

$$f(t,z) \geq \pi_t\sigma_t\theta_t - \frac{1}{2}\eta|\pi_t\sigma_t - z|^2, \quad \forall t, z. \tag{2.3.5}$$

Then, there exists a unique family of decreasing processes $N = \{N^\pi : \pi \in \mathcal{A}\}$ such that

1. *$K_t^\pi = M_t^\pi N_t^\pi$ for all t and π*

2. *K^π is a supermartingale for all π.*

Proof. In order to prove the first claim of the theorem, we explicitly construct the family N. Let $\pi \in \mathcal{A}$, and $t \in [0,T]$. We have

$$K_t^\pi = N_t^\pi M_t^\pi = N_t^\pi \exp(-\eta(x - Y_0)) \exp\left(-\int_0^t \eta(\pi_s\sigma_s - Z_s)\, dB_s \right. \\ \left. -\frac{1}{2}\int_0^t \eta^2(\pi_s\sigma_s - Z_s)^2\, ds\right).$$

This implies

$$N_t^\pi = -\exp\left(\eta(x - Y_0) - \eta(X_t^\pi - Y_t) + \int_0^t \eta(\ \pi_s\sigma_s - Z_s)\, dB_s \right. \\ \left. + \frac{1}{2}\int_0^t \eta^2(\pi_s\sigma_s - Z_s)^2\, ds\right).$$

In this expression, we replace X_t^π (and Y_t) using Equation (1.2.4) (and Equation (2.3.3)). After some cancellations, we are led to

$$N_t^\pi = -\exp\left(\int_0^t \left\{-\eta\pi_s\sigma_s\theta_s + \eta f(s, Z_s) + \frac{1}{2}\eta^2|\pi_s\sigma_s - Z_s|^2\right\} ds\right).$$

Since f satisfies Condition (2.3.5), N^π is a decreasing process.

By Itô's formula, we have

$$\begin{aligned} &K_t^\pi - K_0^\pi \\ &= \int_0^t U'(X_s^\pi - Y_s)(\pi_s\sigma_s - Z_s)\, dB_s + \int_0^t U'(X_s^\pi - Y_s)(\pi_s\sigma_s\theta_s - f(s, Z_s))\, ds \\ &\quad + \frac{1}{2}\int_0^t U''(X_s^\pi - Y_s)(\pi_s\sigma_s - Z_s)^2\, ds \\ &= \int_0^t U'(X_s^\pi - Y_s)(\pi_s\sigma_s - Z_s)\, dB_s - \int_0^t e^{-\eta(X_s^\pi - Y_s)}\big(-\eta\pi_s\sigma_s\theta_s + \eta f(s, Z_s) \\ &\quad + \frac{1}{2}\eta^2(\pi_s\sigma_s - Z_s)^2\big)\, ds. \end{aligned}$$

Because f satisfies Condition (2.3.5), the process $\int_0^t e^{-\eta x}\big(-\eta\pi_s\sigma_s\theta_s+\eta f(s, Z_s)+\frac{1}{2}\eta^2(\pi_s\sigma_s - Z_s)^2\big)\, ds$ is non-decreasing. Hence K^π is the sum of a constant, a martingale and a decreasing process. Therefore, by Doob-Meyer decomposition K^π is a supermartingale. $\square$

An alternative approach to prove the previous result is to use the multiplicative Doob-Meyer decomposition to show that K^π is a local supermartingale, and use the uniform integrability of the stopped process along with the boundedness of Y to conclude. This is the method used in Hu *et al.* (2005), see (the last part of the proof of) Theorem 7 of the aforecited paper.

Now, it remains for us to formally define the function f and to justify the existence and uniqueness of (Y, Z) (note that two different solutions of BSDE (2.3.3) could lead to two value functions of the control problem, which is a contradiction). We construct the function f based on the observation that in Condition (2.3.5), if the equality holds, then $N_t^\pi = -1$ for all t, which implies that $K^\pi = -M^\pi$ is a martingale, and the processes π^* for which this happens are optimal. We have:

$$\begin{aligned}\pi_t\sigma_t\theta_t - \frac{1}{2}\eta|\pi_t\sigma_t - z|^2 &= \pi_t\sigma_t\theta_t - \frac{1}{2}\eta\sigma_t^2|\pi_t|^2 - \frac{1}{2}\eta|z|^2 + \eta\pi_t\sigma_t z \\ &= -\frac{1}{2}\eta\sigma_t^2|\pi_t|^2 - \frac{1}{2}\eta|z|^2 + \pi_t\sigma_t(\theta_t + \eta z) \\ &= -\frac{1}{2}\eta\sigma_t^2\left|\pi_t - \frac{1}{\sigma_t}\left(\frac{\theta_t}{\eta} + z\right)\right|^2 + \frac{1}{2}\eta\left(\frac{\theta_t}{\eta} + z\right)^2 - \frac{1}{2}\eta|z|^2 \\ &= -\frac{1}{2}\eta\sigma_t^2\left|\pi_t - \frac{1}{\sigma_t}\left(\frac{\theta_t}{\eta} + z\right)\right|^2 + \frac{1}{2\eta}\theta_t^2 + \theta_t z. \qquad (2.3.6)\end{aligned}$$

Condition (2.3.5) on f becomes

$$f(t,z) \geq -\frac{1}{2}\eta\sigma_t^2\left|\pi_t - \frac{1}{\sigma_t}\left(\frac{\theta_t}{\eta} + z\right)\right|^2 + \frac{1}{2\eta}\theta_t^2 + \theta_t z.$$

Choose

$$f(t,z) = -\frac{1}{2}\eta\sigma_t^2 dist_t^2\left(\frac{1}{\sigma}\left(\frac{\theta}{\eta} + z\right), C\right) + \frac{1}{2\eta}{}_t^2 + \theta_t z. \qquad (2.3.7)$$

Since $dist\left(\frac{1}{\sigma}\left(\frac{\theta}{\eta} + Z\right), C\right) = \min\left\{\left|\pi - \frac{1}{\sigma}\left(\frac{\theta}{\eta} + Z\right)\right| : \pi \in C\right\}$, Condition (2.3.5) is satisfied for this choice of f.

Remark 2.3.5. *The closeness property of the set C implies that there exists at least one $\pi^* \in C$ realising the minimal distance of $\frac{1}{\sigma}\left(\frac{\theta}{\eta} + Z\right)$ with C. In other words,*

$$\Pi_C\left(\frac{1}{\sigma}\left(\frac{\theta}{\eta} + Z\right)\right) \neq \emptyset,$$

where for a given α, $\Pi_C(\alpha) = \{\beta \in C : |\alpha - \beta| = dist(\alpha, C)\}$. Therefore, $\mathcal{O}pt = \Pi_C\left(\frac{1}{\sigma}\left(\frac{\theta}{\eta} + Z\right)\right) \cap \mathcal{A}$ is the set of optimal policies (Proposition 2.3.10 shows that $\mathcal{O}pt$ is non-empty). If in addition the set C is convex, then there exists exactly one optimal strategy π^*.

The following lemmas will be useful to justify the existence of a process (Y, Z) satisfying Equation (2.3.3) with the function f defined by (2.3.7) as well as to prove that $\mathcal{O}pt \neq \emptyset$.

Lemma 2.3.6. *There exists $k > 0$ such that* $\min\{|a| : a \in C\} \leq k$.

Proof. If a such k does not exist, then for all $l > 0$, $\min\{|a| : a \in C\} > l$. Since C is not empty, by taking $a_1 \in C$, we have $\min\{|a| : a \in C\} > |a_1|$ which is a contradiction. □

Lemma 2.3.7 (Measurable selection)**.** *Let $(a_t)_{t\in[0,T]}$ be a $\mathbb{R}$-valued predictable stochastic process, $C \subset \mathbb{R}$ a closed set.*

1. *The process $d = (dist(a_t, C))_{t\in[0,T]}$ is predictable*
2. *There exists a predictable process a^* with $a^*_t \in \Pi_C(a_t)$, for all $t \in [0,T]$.*

Proof. See Hu *et al.* (2005), Lemma 11. □

Let $z \in \mathbb{R}$ and $t \in [0,T]$,

$$
\begin{aligned}
dist^2\left(\frac{1}{\sigma}\left(\frac{\theta}{\eta}+z\right), C\right) &= \left(\min\left\{\left|\pi - \frac{1}{\sigma}\left(\frac{\theta}{\eta}+z\right)\right| : \pi \in C\right\}_t\right)^2 \\
&\leq \left(\frac{1}{\sigma_t}|z| + \frac{\theta_t}{\eta\sigma_t} + \min\{|\pi| : \pi \in C\}_t\right)^2 \\
&\leq \left(\frac{1}{\sigma_t}|z| + \frac{\theta_t}{\eta\sigma_t} + k\right)^2 \quad \text{(Lemma 2.3.6)} \\
&\leq \frac{1}{\sigma_t^2}|z|^2 + 2\left(\frac{\theta_t}{\eta\sigma_t} + k\right)\frac{1}{\sigma_t}|z| + \frac{\theta_t^2}{\eta^2\sigma_t^2}.
\end{aligned}
$$

Hence, f satisfies

$$
|f(t,z)| \leq k + k_1|z| + k_2|z|^2, \quad k, k_1, k_2 > 0.
$$

Section 3.3.2 addresses the issue of existence of BSDE, but beforehand let us mention that Lemma 2.3.7 implies that $(f(t,z))_{t\in[0,T]}$ is a predictable process, for z fixed. Furthermore, according to Kobylanski (2000), BSDE (2.3.3) has at least one solution $(Y,Z) \in L^\infty(\mathbb{R}) \times H^2(\mathbb{R}^d)$ if f is quadratic and F essentially bounded (we recall that the boundedness property of F was assumed in the settings of the model). The uniqueness follows from Hu *et al.* (2005), Theorem 7, where the authors use the BMO property of the stochastic integral of Z given by the following result.

Proposition 2.3.8. *Let $(Y,Z) \in L^\infty(\mathbb{R}) \times H^2(\mathbb{R}^d)$ be a solution of BSDE* (2.3.3)*, and let $\pi^* = a^*$ constructed as in Lemma 2.3.7 for $a_t = \frac{1}{\sigma_t}\left(\frac{\theta_t}{\eta} + Z_t\right)$. Then the processes*

$$
\int_0^\cdot Z_s\, dB_s \quad \textit{and} \quad \int_0^\cdot \pi^*_s \sigma_s\, dB_s
$$

are BMO-martingales with respect to the probability P.

Proof. See Hu *et al.* (2005), Lemma 12. □

Remark 2.3.9. *The fact that $\int_0^\cdot \pi_s^* \sigma_s \, dB_s$ is a BMO-martingale has the following interesting financial consequence. The trader should only allow strategies yielding an investment with finite credit line.*

Proposition 2.3.10. *We have $\mathcal{O}pt$ is non-empty i.e. $\mathcal{O}pt \neq \emptyset$, and the value function is given by*

$$V^F(x) = -\exp(-\eta(x - Y_0)).$$

Proof. We recall Definition 1.2.3 of an admissible strategy in our setting. Let $\pi^* = a^*$ constructed as in Lemma 2.3.7 for $a_t = \frac{1}{\sigma_t}\left(\frac{\theta_t}{\eta} + Z_t\right)$. Then π^* is predictable. Since $\int_0^\cdot \pi_s^* \sigma_s \, dB_s$ is a BMO-martingale the process $\pi^* \sigma$ is square integrable. It remains to show only that the family $\{U(X_\tau^{\pi^*}) : \tau \text{ stopping time}\}$ is uniformly integrable in order to conclude that π^* is admissible, and thus that $\mathcal{O}pt$ is non-empty. This is also a consequence of the BMO property of $\int_0^\cdot Z_s \, dB_s$ and $\int_0^\cdot \pi_s^* \sigma_s \, ds$. In fact, the process $(M_t^{\pi^*})_{t \in [0,T]}$ (see Equation (2.3.4)) is uniformly integrable thanks to Proposition 2.3.8 and Theorem A.1.2. Since $K^{\pi^*} = -M^{\pi^*}$, we have for all stopping times $\tau \leq T$ $U(X_\tau^{\pi^*}) = -\exp(-\eta Y_\tau) M_\tau^{\pi^*}$. Thus, the boundedness of Y implies that the family $\{U(X_\tau^{\pi^*}) : \tau \text{ stopping time}\}$ is uniformly integrable. Finally, from Equations (2.3.1) and (2.3.2) we have $V^F(x) = -\exp(-\eta(x - Y_0))$. □

In the rest of this section, we will study Problem (1.2.7), still by means of the martingale optimality principle, but now in the case where the investor has a different behaviour towards the risk, i.e. if the investor has a utility function different from $U(x) = -\exp(-\eta x)$.

2.3.1.2 Case of CRRA Utility Functions

In this subsection, we will assess the stochastic control problem (1.2.7) in the case where the investor has a utility U of the class CRRA, i.e. such that the relative risk aversion $-xU''(x)/U'(x)$ is constant. It is well know that this class of utilities can be restricted to the power utility and the logarithm utility. For simplicity[2] we will assume the terminal liability to be zero in both cases. We assume that the investor consumes wealth at a positive rate c_t for all t (i.e. $\alpha \neq 0$) and that the following[3] holds:

(G) The set of constraints C' satisfies: There exist at least one $\tilde{c}^* \in C'$ for which the function $\tilde{c} \mapsto \kappa U(\tilde{c}) - \tilde{c}$, (with $\kappa > 0$) reaches its maximum value.

[2]See the discussion of Section 2.4.

[3]$\tilde{\pi}$ and $\tilde{c}$ are were introduced in Equation (1.2.6).

- Power Utility

The investor has an utility of the form $U(x) = \frac{x^\eta}{\eta}$, $\eta \in (0,1)$. The set $\mathcal{A}$ of admissible strategies is defined by Definition 1.2.3 with the additional requirement $E\left[\int_0^T c_t^\eta \, dt\right] < \infty$.

We would like to construct a family of processes $K = \{K^{\pi,c} = (K_t^{\pi,c})_{t\in[0,T]}\}$ endowed with properties **P1-P4**. From the dynamics of the wealth process given by Equation (1.2.6), we have for all $t \in [0,T]$

$$X_t^{\pi,c} = x\mathcal{E}\left(\int_0^\cdot \tilde{\pi}_s\sigma_s \, dB_s^Q\right)_t \exp\left(-\int_0^t \tilde{c}_s \, ds\right),$$

with $Q = \mathcal{E}\left(-\int_0^\cdot \theta_s \, dB_s\right)_T \cdot P$. Thus,

$$U(X_t^{\pi,c}) = \frac{x^\eta}{\eta}\exp\left(\int_0^t \eta\tilde{\pi}_s\sigma_s \, dB_s^Q - \frac{1}{2}\int_0^t \eta\tilde{\pi}_s^2\sigma_s^2 \, ds\right)\exp\left(-\int_0^t \eta\tilde{c}_s \, ds\right).$$

From Property **P1**, we put for all $\pi \in \mathcal{A}$ and $t \in [0,T]$

$$\begin{aligned} K_t^{\pi,c} &= \frac{x^\eta}{\eta}\exp\left(\int_0^t \eta\tilde{\pi}_s\sigma_s \, dB_s^Q - \frac{1}{2}\int_0^t \eta\tilde{\pi}_s^2\sigma_s^2 \, ds + Y_t\right)\exp\left(-\int_0^t \eta\tilde{c}_s \, ds\right) \\ &\quad + \int_0^t \alpha\frac{1}{\eta}c_s^\eta ds \\ &= U(X_t^{\pi,c})e^{Y_t} + \int_0^t \alpha U(c_s) \, ds, \end{aligned}$$

where (Y,Z) is solution of the BSDE

$$Y_t = 0 - \int_t^T Z_s \, dB_s - \int_t^T f(s, Y_s, Z_s) \, ds. \tag{2.3.8}$$

Proposition 2.3.11. *Assume that the generator f is such that (Y,Z) exists, and that $f(t,y,z)$ satisfies the condition*

$$f(t,y,z) \le -\eta\tilde{\pi}_t\sigma_t(\theta_t + z) - \eta\frac{\eta-1}{2}\tilde{\pi}_t^2\sigma_t^2 - \frac{1}{2}z^2 - (\alpha\tilde{c}_t^\eta e^{-y} - \eta\tilde{c}_t). \tag{2.3.9}$$

Then $K^{\pi,c}$ is a supermartingale (with respect to the probability P) for all $(\tilde{\pi}, \tilde{c})$ and if the equality holds then $K^{\pi,c}$ is a martingale.

Proof. Let $t \in [0,T]$ and $(\pi, c) \in \mathcal{A}$. Itô's formula yields

$$\begin{aligned} &K_t^{\pi,c} - K_0^{\pi,c} \\ &= \int_0^t (X_s^{\pi,c})^\eta e^{Y_s}(\tilde{\pi}_s\sigma_s + \frac{1}{\eta}Z_s) \, dB_s + \int_0^t (X_s^{\pi,c})^\eta e^{Y_s}\left(\tilde{\pi}_s\sigma_s\theta_s + \frac{\eta-1}{2}\,\tilde{\pi}_s^2\sigma_s^2\right. \\ &\quad \left. + \frac{1}{\eta}f(s, Y_s, Z_s) + \frac{1}{2\eta}Z_s^2 + Z_s\tilde{\pi}_s\sigma_s - \tilde{c}_s\right) ds + \int_0^t \alpha\frac{1}{\eta}c_s^\eta \, ds. \end{aligned}$$

A simple rearrangement leads to

$$\begin{aligned} &K_t^{\pi,c} - K_0^{\pi,c} \\ &= \int_0^t (X_s^{\pi,c})^\eta e^{Y_s}(\tilde{\pi}_s\sigma_s + \frac{1}{\eta}Z_s)\,dB_s \\ &\quad + \int_0^t (X_s^{\pi,c})^\eta e^{Y_s}\left(\tilde{\pi}_s\sigma_s(\theta_s + Z_s) + \frac{\eta-1}{2}\,\tilde{\pi}_s^2\sigma_s^2 + \frac{1}{2\eta}Z_s^2 + \frac{1}{\eta}f(s,Y_s,Z_s)\right)ds \\ &\quad + \int_0^t (X_s^{\pi,c})^\eta e^{Y_s}(\alpha\frac{1}{\eta}\tilde{c}_s^\eta e^{-Y_s} - \tilde{c}_s)\,ds. \end{aligned} \tag{2.3.10}$$

The first term in (2.3.10) is a martingale, and since f satisfies Condition (2.3.9) the last two terms of (2.3.10) form a decreasing process with integrable total variation. Therefore, $K^{\pi,c}$ is a supermartingale. If the equality holds in (2.3.9), then the last two terms of (2.3.10) vanish and K^π becomes a martingale.

□

Now let us construct the generator f. One can write

$$\tilde{\pi}_s\sigma_s(\theta_s + Z_s) + \frac{\eta-1}{2}\tilde{\pi}_s^2\sigma_s^2 = \sigma_s^2\frac{\eta-1}{2}\left|\tilde{\pi}_s - \frac{Z_s+\theta_s}{\sigma_s(1-\eta)}\right|^2 + \frac{1}{2(1-\eta)}|Z_s+\theta_s|^2.$$

Condition (2.3.9) becomes

$$f(t,y,z) \le -\eta\sigma_t^2\frac{\eta-1}{2}\left|\tilde{\pi}_t - \frac{Z_t+\theta_t}{\sigma_t(1-\eta)}\right|^2 - \frac{\eta}{2(1-\eta)}|z+\theta_t|^2 - \frac{1}{2}|z|^2 - (\alpha\tilde{c}_t^\eta e^{-y} - \eta\tilde{c}_t).$$

Let us choose

$$\begin{aligned} &f(t,y,z) = \\ &\eta\sigma_t^2\frac{1-\eta}{2}dist_t^2\left(\frac{z+\theta}{\sigma(1-\eta)}, C\right) - \frac{\eta}{2(1-\eta)}|z+\theta_t|^2 - \frac{1}{2}|z|^2 - \max_{\tilde{c}\in C'}(\alpha\tilde{c}^\eta e^{-y} - \eta\tilde{c}). \end{aligned} \tag{2.3.11}$$

For this choice of f, Condition (2.3.9) is satisfied, and the equality holds if and only if $\tilde{\pi} \in \Pi_C\left(\frac{z+\theta}{\sigma(1-\eta)}\right)$ and $\tilde{c}$ is such that $\max_{\tilde{c}\in C'}(\alpha\tilde{c}^\eta e^{-y} - \eta\tilde{c})$ is attained at $\tilde{c}$, i.e.

$$\tilde{c} \in \arg\max_{\tilde{c}\in C'}(\alpha\tilde{c}^\eta e^{-y} - \eta\tilde{c}).$$

Note that $\Pi_C\left(\frac{z+\theta}{\sigma(1-\eta)}\right)$ is non-empty since C is closed. From Assumption **(G)**, the set $\arg\max_{\tilde{c}\in C'}(\alpha\tilde{c}^\eta e^{-y} - \eta\tilde{c})$ is non-empty. Hence, for

$$(\tilde{\pi}^*, \tilde{c}^*) \in \Pi_C\left(\frac{z+\theta}{\sigma(1-\eta)}\right) \times \arg\max_{\tilde{c}\in C'}(\alpha\tilde{c}^\eta e^{-y} - \eta\tilde{c}), \tag{2.3.12}$$

the process K^{π^*,c^*} is a martingale.

Moreover, due to Lemma 2.3.7 and assumption **(G)**, $(f(t,y,z))_{t\in[0,T]}$ is a predictable process for fixed y and z. In addition f is of quadratic growth in z and, up to a change of variable, of linear growth in y. Hence, according to Kobylanski (2000), BSDE (2.3.8) has a unique solution such that Y is essentially bounded.

It remains for us to show that there exist $(\tilde{\pi}^*, \tilde{c}^*)$ satisfying (2.3.12) which are admissible. This is a consequence of Proposition 2.3.8 and assumption **(G)**. Let $\tilde{\pi}^*$ constructed like in Lemma 2.3.7 for $a = \frac{z+\theta}{\sigma(1-\eta)}$. Then, $\tilde{\pi}^*$ is predictable. Let $\tau < T$ be a stopping time. The wealth process is given by $X_\tau^{\pi^*,c^*} = x\mathcal{E}\left(\int_0^\cdot \tilde{\pi}_s^*\sigma_s\,dB_s^Q\right)_\tau \exp\left(-\int_0^t \tilde{c}_s^*\,ds\right)$. Since $\int_0^\cdot \tilde{\pi}_s^*\sigma_s\,dB_s$ is a BMO-martingale with respect to the probability P and because θ is bounded, it is also a BMO-martingale with respect to the probability Q, see Theorem A.1.2. In addition, $\tilde{c}^*$ and F are bounded. Hence we conclude that $X_\tau^{\pi^*,c^*}$ is uniformly integrable. The strategy $(\tilde{\pi}^*, \tilde{c}^*)$ is therefore admissible and by the martingale optimality principle, it is an optimal strategy. Finally, from Equations (2.3.1) and (2.3.2) the value function is given by

$$V^F(x) = \frac{1}{\eta}x^\eta \exp(Y_0).$$

Let us turn to the case where the investor has a logarithmic utility.

- Logarithmic Utility

We assume that the utility of the investor is given by $U(x) = \log(x)$. The set $\mathcal{A}$ of admissible strategies is still defined by Definition 1.2.3 with the additional requirement

$$E\left[\int_0^T |\log(\tilde{c}_t)|\,dt + \int_0^T \tilde{c}_t\,dt\right] < \infty.$$

For every admissible pair $(\tilde{\pi}, \tilde{c})$ and $t \in [0,T]$, put

$$K_t^{\pi,c} = h(t)(\log(X_t^{\pi,c}) - Y_t) + \int_0^t \alpha \log(\tilde{c}_s)\,ds;$$

where $h(t) = T + 1 - t$ and (Y,Z) is solution of the BSDE

$$Y_t = 0 - \int_t^T Z_s\,dB_s - \int_t^T f(s,Y_s)\,ds. \tag{2.3.13}$$

The following result holds:

Proposition 2.3.12. *Assume that the generator f is such that (Y,Z) exists and that $f(t,y)$ satisfies the condition*

$$f(t,y) \geq \tilde{\pi}_t\sigma_t\theta_t - \tilde{c}_t - \frac{1}{2}\tilde{\pi}_t^2\sigma_t^2 + \frac{\alpha\log(\tilde{c}_t) + y}{h(t)}. \tag{2.3.14}$$

Then $K_t^{\pi,c}$ is a supermartingale for all (π,c) and if the equality holds then $K^{\pi,c}$ is a martingale.

Proof. Let $t \in [0,T]$ and $\pi \in \mathcal{A}$. Itô's formula yields

$$K_t^{\pi,c} - K_0^{\pi,c} = \int_0^t h(s)(\tilde{\pi}_s\sigma_s - Z_s)\,dB_s + \int_0^t \Big(\tilde{\pi}_s\sigma_s\theta_s - \tilde{c}_s - \frac{1}{2}\tilde{\pi}_s^2\sigma_s^2 \quad (2.3.15)$$

$$+ \frac{\alpha\log(\tilde{c}_s) + Y_s}{h(s)} - f(s, Y_s)\Big)\,ds. \quad (2.3.16)$$

Since f satisfies Condition (2.3.14), the finite variation part of the above process is decreasing. Hence, $K^{\pi,c}$ is the sum of a constant, a martingale and a decreasing process with integrable total variation. Therefore, $K^{\pi,c}$ is a supermartingale. If the inequality in Condition (2.3.14) becomes an equality the second term in (2.3.15) vanishes. Hence, $K^{\pi,c}$ is a martingale. □

Observing that

$$\tilde{\pi}_t\sigma_t\theta_t - \frac{1}{2}\tilde{\pi}_t^2\sigma_t^2 = -\frac{\sigma_t^2}{2}\left|\tilde{\pi}_t - \frac{\theta_t}{\sigma_t}\right|^2 + \frac{1}{2}\theta_t^2,$$

we choose

$$f(t,y) = -\frac{\sigma_t^2}{2}dist_t^2(\frac{\theta}{\sigma}, C) + \max_{\tilde{c}\in C'}(\alpha\frac{\log(\tilde{c})}{h} - \tilde{c}) + \frac{y}{h(t)} + \frac{1}{2}\theta_t^2.$$

For this choice of f, Condition (2.3.14) is satisfied, and the equality holds if and only if $\tilde{\pi} \in \Pi\left(\frac{\theta}{\sigma}\right)$, which is non-empty since C is closed, and $\tilde{c} \in \arg\max_{\tilde{c}\in C'}(\alpha\frac{\log(\tilde{c})}{h} - \tilde{c})$ which is non-empty from Assumption **(G)**. Hence, K^{π^*,c^*} is a martingale for all

$$(\tilde{\pi}^*, \tilde{c}^*) \in \Pi_C\left(\frac{\theta}{\sigma}\right) \times \arg\max_{\tilde{c}\in C'}(\alpha\frac{\log(\tilde{c})}{h} - \tilde{c}). \quad (2.3.17)$$

Note that f is of linear growth in y and does not depend on z. Thus, it follows from El Karoui *et al.* (1997) that (2.3.13) has a unique solution. There exist pairs of processes $(\tilde{\pi}^*, \tilde{c}^*)$ satisfying (2.3.17) which are admissible. In fact, Let $\tilde{\pi}^*$ be constructed like in Lemma 2.3.7 for $a = \frac{\theta}{\sigma}$. Then $\tilde{\pi}^*$ is predictable. Since $\int_0^\cdot \sigma_s\tilde{\pi}_s^*\,dB_s$ and $\int_0^\cdot Z_s\,dB_s$ are BMO-martingales,

$$K_t^{\pi^*,c^*} = K_0^{\pi^*,c^*} + \int_0^t h(s)(\sigma_s\tilde{\pi}_s^* - Z_s)\,dB_s$$

is uniformly integrable (apply Itô isometry and use Cauchy-Schwarz inequality since h is square integrable). In addition, $c^* \in \arg\max_{\tilde{c}\in C'}(\alpha\frac{\log(\tilde{c})}{h} - \tilde{c})$ implies that $\alpha\frac{\log(\tilde{c}^*)}{h} - \tilde{c}^*$ is bounded. For all $t \in [0,T]$, $\tilde{c}_t \mapsto \log(\tilde{c}_t)$ is defined on $\mathbb{R}$. It is easy to see that $\alpha\frac{\log(\tilde{c}_t)}{h(t)} - \tilde{c}_t$ reaches its maximum at $\tilde{c}_t^* = \alpha/h(t)$. Hence $\log(\tilde{c}^*)$ and $\tilde{c}^*$ are bounded. Therefore,

$$\log(X_\tau^{\pi^*,c^*}) = \frac{1}{h}\left(K_\tau^{\pi^*,c^*} - \int_0^\tau \alpha\log(\tilde{c}_s^*)\,ds\right) + Y_\tau$$

is uniformly integrable and

$$E\left[\int_0^T |\log(\tilde{c}_t^*)|\,dt + \int_0^T \tilde{c}_t^*\,dt\right] < \infty.$$

Finally, from Equations (2.3.1) and (2.3.2) the value function is given by

$$V^F(x) = (T+1)(\log(x) - Y_0).$$

The method discussed in the next subsection is, in some ways, a generalization of the method that we have described in this subsection. We will use the stronger optimality criteria given by Corollary 2.2.3 and Proposition 2.2.4. Note that Equation (2.3.23) is valid for general utility functions.

2.3.2 Characterization via Itô-Ventzell's Formula

We consider the stochastic control problem (1.2.7) with the trader's utility function $U(x) = -e^{-\eta x}$, $\eta \in (0,1)$. Put $Y^{\pi,c}(s,x) = V^F(s, X_s^{\pi,c})$ and $Y(s,x) = V^F(s,x)$, for $(\pi,c) \in \mathcal{A}$, $s \in [t,T]$. In this subsection, based on a reasoning of Mania and Tevzadze (2008), we will show that $Y^{\pi^*,c^*}(\cdot,x)$ can be expressed in terms of the first component of the solution of a BSDE for (π^*,c^*) optimal. The BSDE that we derive was first obtained by Mania and Tevzadze (2008) in the fairly general case where the price process is a continuous semimartingale. Musiela and Zariphopoulou (2010) also obtained the same equation in a Brownian setting and with the assumption that the value function is an Itô-diffusion. Note that neither of these works include consumption in their model. Thus let $(\pi,c) \in \mathcal{A}$. From Proposition 2.2.4, $Y^{\pi,c}(\cdot,x)$ is a supermartigale. Moreover, by admissibility of (π,c), $Y^{\pi,c}(\cdot,x)$ is of class D. Hence, by Doob-Meyer decomposition there exist two processes $A(\cdot,x)$, $M(\cdot,x)$ such that for all $s \in [t,T]$

$$Y^{\pi,c}(s,x) = Y^{\pi,c}(t,x) + A(s, X_s^{\pi,c}) + M(s, X_s^{\pi,c}) \tag{2.3.18}$$

with $M(\cdot,x)$ a martingale and $A(\cdot,x)$ a decreasing process with finite variation. We assume that there exists $a(\cdot,\cdot)$ such that $A(s, X_s^{\pi,c}) = \int_t^s a(u, X_u^{\pi,c})\,du$. Besides, by the martingale representation theorem, there exists $Z(\cdot,x)$ such that $M(s, X_s^{\pi,c}) = M(t,x) + \int_t^s Z(u, X_u^{\pi,c})\,dB_u$, $s \geq t$. Since V^F is a random function (it is written as a conditional expectation) Itô's formula cannot be applied. Rather, we will use Itô-Ventzell's formula, an extension of Itô's formula to random functions stated as follows.

Proposition 2.3.13. *Assume that V^F is strictly concave, the mapping $s \mapsto Y^{\pi,c}(s,x)$ is twice continuously differentiable for all (ω,t) with the first derivative $Y^{\pi,c}(\cdot,x)$ satisfying*

$$Y_x^{\pi,c}(s,x) = Y_x^{\pi,c}(t,x) + \widehat{A}(s, X_s^{\pi,c}) + \int_t^s Z_x(u, X_u^{\pi,c})\,du + M_x(s,x)$$

and $Y^{\pi,c}_{xx}(\cdot,x)$ *is RCLL for every* $x \in \mathbb{R}$. *Then,* $Y^{\pi,c}(s,x) = Y^{\pi,c}(t,x) + A(s,X^{\pi,c}_s) + M(s,X^{\pi,c}_s)$ *with*

$$\begin{aligned} &A(s,X^{\pi,c}_s) - A(t,x) \\ &= \int_t^s \left[Y^{\pi,c}_x(u,x)\,dX^{\pi,c}_u + \frac{1}{2}Y^{\pi,c}_{xx}(u,x)\,d[X^{\pi,c}]_u + Z_x(u,X^{\pi,c}_u)\,d[B,X^{\pi,c}]_u \right] \\ &\qquad + \int_t^s a(u,X^{\pi,c}_u)\,du. \end{aligned}$$

Proof. See Mania and Tevzadze (2008), Proposition 2.2 and the references therein. □

Recall that the wealth process is defined by Equation (1.2.5). Using Itô-Ventzell's formula, we have for all $(\pi,c) \in \mathcal{A}$, $t \le s \le T$ and $x \in \mathbb{R}_+$

$$\begin{aligned} &Y^{\pi,c}(s,x) \\ &= Y^{\pi,c}(t,x) + M(s,X^{\pi,c}_s) + A(t,x) + \int_t^s \big[\pi_u\sigma_u\left(\theta_u Y^{\pi,c}_x(u,x) + Z_x(u,X^{\pi,c}_u)\right) \\ &+ \frac{1}{2}Y^{\pi,c}_{xx}(u,x)\sigma_u^2\pi_u^2 - Y^{\pi,c}_x(u,x)c_u\big]\,du + \int_t^s Y^{\pi,c}_x(u,x)\sigma_u\pi_u\,dB_u \\ &+ \int_t^s a(u,X^{\pi,c}_u)\,du \\ &= Y^{\pi,c}(t,x) + M(s,x) + \int_t^s a(u,X^{\pi,c}_u)\,du + \int_t^s \big[\frac{1}{2}Y^{\pi,c}_{xx}(u,x)\left(g(u,X^{\pi,c}_u) + \pi_u\sigma_u\right)^2 \\ &\quad - \frac{\left|(\theta_u - \frac{c_u}{\pi_u\sigma_u})Y^{\pi,c}_x(u,x) + Z_x(u,X^{\pi,c}_u)\right|^2}{2Y^{\pi,c}_{xx}(u,x)}\big]\,du + \int_t^s Y^{\pi,c}_x(u,x)\sigma_u\pi_u\,dB_u + A(t,x), \end{aligned} \tag{2.3.19}$$

with $g(s,x) = \frac{(\theta_s - \frac{c_s}{\pi_s\sigma_s})Y^{\pi,c}_x(s,x) + Z_x(s,x)}{Y^{\pi,c}_{xx}(s,x)}$. Since $Y^{\pi,c}(\cdot,x)$ is a supermartingale, the finite variation part of the above decomposition is a decreasing process. That is, for all $(\pi,c) \in \mathcal{A}$

$$a(s,X^{\pi,c}_s) \le \frac{\left|(\theta_s - \frac{c_s}{\pi_s\sigma_s})Y^{\pi,c}_x(s,x) + Z_x(s,x)\right|^2}{2Y^{\pi,c}_{xx}(s,x)} - \frac{1}{2}Y^{\pi,c}_{xx}(s,x)\big(g(s,X^{\pi,c}_s) + \pi_s\sigma_s\big)^2.$$

This implies

$$a(t,x) \le \frac{\left|(\theta_t - \frac{c_t}{\pi_t\sigma_t})Y^{\pi,c}_x(t,x) + Z_x(t,x)\right|^2}{2Y^{\pi,c}_{xx}(t,x)} - \frac{1}{2}Y^{\pi,c}_{xx}(t,x)\big(g(t,x) + \pi_t\sigma_t\big)^2.$$

That is,

$$a(t,x) \leq \operatorname*{ess\ inf}_{(\pi,c)\in\mathcal{A}} \frac{\left|(\theta_t - \frac{c_t}{\pi_t\sigma_t})Y_x^{\pi,c}(t,x) + Z_x(t,x)\right|^2}{2Y_{xx}^{\pi,c}(t,x)} + \operatorname*{ess\ inf}_{(\pi,c)\in\mathcal{A}} \left\{-\frac{1}{2}Y_{xx}^{\pi,c}(t,x)\left(g(t,x)+\pi_t\sigma_t\right)^2\right\}.$$

Since $Y_{xx} < 0$ it follows from Mania and Tevzadze (2008) Lemma A.1 that the last term of the right hand side in the latter inequality is zero. Thus,

$$a(t,x) \leq \frac{\left|(\theta_t - \frac{c_t}{\pi_t\sigma_t})Y_x^{\pi,c}(t,x) + Z_x(t,x)\right|^2}{2Y_{xx}^{\pi,c}(t,x)} \quad \lambda \text{ a.s. for all } (\pi,c)\in\mathcal{A}. \tag{2.3.20}$$

This inequality will provide a key argument in the proof of the following result.

Theorem 2.3.14. *Under assumptions of Proposition 2.3.13, the value function is given by $V^F(t,x) = e^{-\eta x}\widehat{Y}_t$ where the process $\widehat{Y}$ is the first component of the solution of the BSDE*

$$\widehat{Y}_t = -e^{\eta F} - \int_t^T \frac{|(\theta_u - \frac{c_u^*}{\pi_u^*\sigma_u})\widehat{Y}_u + \widehat{Z}_u|^2}{\widehat{Y}_u}\,du - \int_t^T \widehat{Z}_u\,dB_u - \widehat{M}_t. \tag{2.3.21}$$

Where (π^*, c^*) is an optimal strategy and $\widehat{M_t} = e^{\eta x}\big(M(t,x) + A(t,x)\big)$.

Proof. Let (π^*, c^*) be an optimal strategy. By Proposition 2.2.4, $Y^{\pi^*,c^*}(\cdot,x)$ is a martingale. Consequently, the finite variation part in the decomposition (2.3.19) is zero, and using Inequality (2.3.20) we are led to

$$Y_{xx}^{\pi^*,c^*}(s,x)\left[\frac{(\theta_s - \frac{c_s^*}{\pi_s^*\sigma_s})Y_x^{\pi^*,c^*}(s,x) + Z_x(s,x)}{Y_{xx}^{\pi^*,c^*}(s,x)} + \pi_s^*\sigma_s\right]^2 \geq 0, \quad \lambda \text{ a.s.}$$

Moreover, $Y_{xx}^{\pi^*,c^*} < 0$. Hence,

$$\frac{(\theta_s - \frac{c_s^*}{\pi_s^*\sigma_s})Y_x^{\pi^*,c^*}(s,x) + Z_x(s,x)}{Y_{xx}^{\pi^*,c^*}(s,x)} + \pi_s^*\sigma_s = 0 \quad \lambda \text{ a.s.}, \tag{2.3.22}$$

which implies

$$a(s, X_s^{\pi^*,c^*}) = \frac{|(\theta_s - \frac{c_s^*}{\pi_s^*\sigma_s})Y_x^{\pi^*,c^*}(s,x) + Z_x(s,x)|^2}{2Y_{xx}^{\pi^*,c^*}(s,x)}.$$

In particular,

$$a(t,x) = \frac{|(\theta_t - \frac{c_t^*}{\pi_t^*\sigma_t})Y_x^{\pi^*,c^*}(t,x) + Z_x(t,x)|^2}{2Y_{xx}^{\pi^*,c^*}(t,x)}.$$

Plugging this in Equation (2.3.18) yields

$$Y(t,x) = Y(T,x) - \int_t^T \frac{|(\theta_u - \frac{c_u^*}{\pi_u^* \sigma_u})Y_x^{\pi^*,c^*}(u,x) + Z_x(u,x)|^2}{2Y_{xx}^{\pi^*,c^*}(u,x)}\,du \\ - \int_t^T Z(u,x)\,dB_u - M(t,x) - A(t,x). \tag{2.3.23}$$

On the other hand, using a property of the exponential function we have

$$Y^{\pi^*,c^*}(t,x) = V^F(t,x) = e^{-\eta x}\widehat{Y}_t$$

with

$$\widehat{Y}_t = \sup_{(\pi,c)\in\mathcal{A}} E\left[-\exp\left\{-\eta\left(\int_t^T \pi_u \frac{dS_u}{S_u} - \int_t^T c_u\,du - F\right)\right\} \mid \mathcal{F}_t\right],$$

where we omit the dependence of $\widehat{Y}$ in (π^*, c^*). Equation (2.3.23) becomes

$$\widehat{Y}_t = -e^{\eta F} - \int_t^T \frac{|(\theta_u - \frac{c_u^*}{\pi_u^* \sigma_u})\widehat{Y}_u - \frac{e^{\eta x}}{\eta} Z_x(u,x)|^2}{\widehat{Y}_u}\,du - \int_t^T e^{\eta x} Z(u,x)\,dB_u \\ - e^{\eta x}\big(M(t,x) + A(t,x)\big).$$

Since $\widehat{Y}$ does not depend on the parameter x, there exists a process $\widehat{M}$ such that $\widehat{M}_s = e^{\eta x}\big(M(s,x) + A(s,x)\big)$ for all x and s. We put $\widehat{Z} = e^{\eta x} Z(\cdot, x)$, then $(\widehat{Y}, \widehat{Z})$ solves the BSDE

$$\widehat{Y}_t = -e^{\eta F} - \int_t^T \frac{|(\theta_u - \frac{c_u^*}{\pi_u^* \sigma_u})\widehat{Y}_u + \widehat{Z}_u|^2}{\widehat{Y}_u}\,du - \int_t^T \widehat{Z}_u\,dB_u - \widehat{M}_t. \tag{2.3.24}$$

□

Remark 2.3.15. *The backward stochastic partial differential equation* (2.3.23) *describes the value function of the control problem for general utility functions. As mentioned by Mania and Tevzadze (2008), it is a verification tool, for it is obtained upon assumptions about the value function, not the market parameters. In addition, there are no existence and/or uniqueness results for that equation.*

In the particular case of exponential utility, Equation (2.3.23) becomes Equation (2.3.21) and the value function $V^F(0,x)$ is known to be twice differentiable. Equation (2.3.21) has a particular generator, that does not belong to the classical classes of Lipschitz or quadratic generators. Imkeller *et al.* (2011) prove existence and uniqueness of solutions of this equation when the terminal condition is bounded away from zero by transforming it, thanks to an exponential change of variable, into a quadratic BSDE. They show that in our

Brownian setting it can even be transformed into a linear BSDE. In the case where the consumption is taken to be zero at all time, i.e. $c = 0$ the optimal strategy is merely given by (2.3.22).

Let us now present a BSDE characterisation which does not rely on any of the optimality principles described above.

2.3.3 Characterization via Martingale Representation

In this subsection, we present another characterization of the value function of the expected exponential utility maximization problem which is used in Sung and Wan (2010) to represent the principal's and the agent's utility in a principal-agent problem. The method does not use the martingale optimality principle. Instead, it is based on the martingale representation theorem and a comparison result for BSDEs.

Consider the problem

$$V(t,x) = \sup_{\pi \in \mathcal{A}} E\left[- \exp\left\{-\eta(X_T^\pi - F)\right\} | \mathcal{F}_t\right] \tag{2.3.25}$$

with X_T^π given by (1.2.4). For all $\pi \in \mathcal{A}$, put

$$\begin{aligned} - \exp\left\{-\eta L_t^\pi\right\} &= E\left[- \exp\left\{-\eta\left(x + \int_t^T \pi_s\sigma_s(\theta_s\, ds + dB_s) - F\right)\right\} \mid \mathcal{F}_t\right] \\ &= E\left[- \exp\left\{-\eta\left(x + \int_t^T \pi_s\sigma_s dB_s^Q - F\right)\right\} \mid \mathcal{F}_t\right], \end{aligned} \tag{2.3.26}$$

where Q is the probability with density $\frac{dQ}{dP} = \mathcal{E}\left(-\int_0^\cdot \theta_s dB_s\right)_T$ and (by Girsanov's theorem[4]) $B^Q = \int_0^\cdot \theta_s\, ds + B$ is a Q-Brownian motion. Next, define

$$J_t = - \exp\left\{-\eta\left(x + \int_0^t \pi_s\sigma_s\, dB_s^Q + L_t^\pi\right)\right\}.$$

Then, $(J_t)_{t\in[0,T]}$ is a $(P-)$ martingale. In fact, for all t

$$\begin{aligned} &J_t \\ &= - \exp\left\{-\eta L_t^\pi\right\} \exp\left\{-\eta\left(x + \int_0^t \pi_s\sigma_s\, dB_s^Q\right)\right\} \\ &= E\left[- \exp\left\{-\eta\left(x + \int_t^T \pi_s\sigma_s\, dB_s^Q - F\right)\right\} \mid \mathcal{F}_t\right] \exp\left\{-\eta\left(x + \int_0^t \pi_s\sigma_s\, dB_s^Q\right)\right\} \\ &= E\left[- \exp\left\{-\eta\left(2x + \int_0^T \pi_s\sigma_s\, dB_s^Q - F\right)\right\} \mid \mathcal{F}_t\right], \end{aligned}$$

[4]Note that $(\theta_t)_{t\in[0,T]}$ is assumed to be uniformly bounded.

which is a martingale because $-\exp\left\{-\eta\left(2x+\int_0^T \pi_s\sigma_s\, dB_s^Q - F\right)\right\}$ is an integrable and $\mathcal{F}_T$-measurable random variable.

Thus, by the martingale representation theorem, there exists a unique predictable square-integrable process $(\widehat{Z}_t)_{t\in[0,T]}$ such that

$$J_t = -\exp\left\{-\eta\left(2x+F+\int_0^T \pi_s\sigma_s\, dB_s^Q\right)\right\} + \int_t^T \widehat{Z}_s\, dB_s.$$

Put $\widehat{Z}_t = J_t\tilde{Z}_t$ for all t. Then,

$$J_t = -\exp\left\{-\eta\left(2x+F+\int_0^T \pi_s\sigma_s\, dB_s^Q\right)\right\} + \int_t^T J_s\tilde{Z}_s\, dB_s. \tag{2.3.27}$$

On the first hand, we differentiate J_t using its expression given by (2.3.27). This gives

$$dJ_t = -\tilde{Z}_t J_t\, dB_t. \tag{2.3.28}$$

On the other hand, we differentiate J_t using its definition and Itô's formula. This leads to

$$dJ_t = -\eta J_t\left(\pi_t\sigma_t\, dB_t^Q + dL_t^\pi\right) + \frac{\eta^2}{2}J_t\left(d[L^\pi,L^\pi]_t + 2\pi_t\sigma_t d[L^\pi,B^Q]_t + \pi_t^2\sigma_t^2\, dt\right). \tag{2.3.29}$$

From Equations (2.3.28) and (2.3.29), we have

$$\pi_t\sigma_t\, dB_t^Q + dL_t^\pi - \frac{\eta}{2}\left(d[L^\pi,L^\pi]_t + 2\pi_t\sigma_t\, d[L^\pi,B^Q]_t + \pi_t^2\sigma_t^2\, dt\right) = \frac{1}{\eta}\tilde{Z}_t\, dB_t.$$

Hence,

$$\begin{aligned} dL_t^\pi &= \frac{1}{2}\eta\pi_t^2\sigma_t^2\, dt - \pi_t\sigma_t\theta_t\, dt \\ &\quad + \left(\frac{1}{\eta}\tilde{Z}_t - \pi_t\sigma_t\right) dB_t + \frac{\eta}{2}\left(d[L^\pi,L^\pi]_t + 2\pi_t\sigma_t\, d[L^\pi,B^Q]_t\right). \end{aligned}$$

Thus, we have the quadratic (and cross) variations

$$d[L^\pi,L^\pi]_t = \left(\frac{1}{\eta}\tilde{Z}_t - \pi_t\sigma_t\right)^2 dt \quad \text{and} \quad \pi_t\sigma_t\, d[L^\pi,B^Q]_t = \left(\frac{1}{\eta}\tilde{Z}_t - \pi_t\sigma_t\right)\pi_t\sigma_t\, dt.$$

Therefore, $dL_t^\pi = \left(\frac{1}{2}\eta\pi_t^2\sigma_t^2 + \frac{\eta}{2}Z_t^2 + \eta Z_t\pi_t\sigma_t - \pi_t\sigma_t\theta_t\right) dt + Z_t\, dB_t$ where we put $Z_t = \frac{1}{\eta}\tilde{Z}_t - \pi_t\sigma_t$ for all t. Hence, we have that for all $\pi\in\mathcal{A}$ the pair of processes (L^π, Z) solves the BSDE

$$L_t^\pi = L_T^\pi + \int_t^T f^\pi(s,Z_s)\, ds - \int_t^T Z_s\, dB_s \tag{2.3.30}$$

with the generator $f^\pi(t,z) = -\frac{1}{2}\eta\pi_t^2\sigma_t^2 + \pi_t\sigma_t(\theta_t - \eta z) - \frac{1}{2}z^2$ and the terminal condition $L_T^\pi = x - F$. Note that from Kobylanski (2000), for each $\pi \in \mathcal{A}$, (L^π, Z) exists and is unique. Observe that for a given π and for all t and z,

$$f^\pi(t,z) = -\frac{1}{2}\eta\sigma_t^2\left(\pi_t - \frac{1}{\eta\sigma_t}(\theta_t - \eta z)\right)^2 + \frac{1}{2\eta}(\theta_t - \eta z)^2 - \frac{1}{2}z^2$$

and the following result holds.

Proposition 2.3.16. *The value function of the stochastic control problem* (2.3.25) *is given by* $V^F(t,x) = -\exp(-\eta L_t)$ *where* L *is the first component of the solution of the BSDE*

$$L_t = x - F + \int_t^T g(s, Z_s)\,ds - \int_t^T Z_s\,dB_s$$

with

$$g(t,z) = -\frac{1}{2}\eta\sigma_t^2 dist_t^2\left(\frac{1}{\eta\sigma}(\theta - \eta z), C\right) + \frac{1}{2\eta}(\theta_t - \eta z)^2 - \frac{1}{2}z^2.$$

Proof. For all t, z put

$$\begin{aligned} g(t,z) &= \sup_{\pi\in C} f^\pi(t,z) \\ &= \sup_{\pi\in C}\left\{-\frac{1}{2}\eta\sigma^2\left|\pi - \frac{1}{\eta\sigma}(\theta - \eta z)\right|^2\right\}_t + \frac{1}{2\eta}(\theta_t - \eta z)^2 - \frac{1}{2}z^2 \\ &= -\frac{1}{2}\eta\sigma_t^2\inf_{\pi\in C}\left|\pi - \frac{1}{\eta\sigma}(\theta - \eta z)\right|_t^2 + \frac{1}{2\eta}(\theta_t - \eta z)^2 - \frac{1}{2}z^2. \end{aligned}$$

The above defined function g is concave and with quadratic growth in z, and the $\mathcal{F}_T$-random variable $x - F$ is bounded. By Kobylanski (2000) , the BSDE

$$L_t = x - F + \int_t^T g(s, Z_s)\,ds - \int_t^T Z_s\,dB_s$$

has a unique solution (L, Z). Moreover, by the comparison result for BSDEs, given by Theorem 3.3.9, $L_t = \sup_{\pi\in C} L_t^\pi$ for all t.

By definition of $(L_t^\pi)_{t\in[0,T]}$ (see Equation (2.3.26)) we have

$$V^F(t,x) = \sup_{\pi\in C}\{-\exp(-\eta L_t^\pi)\}.$$

Since the risk aversion η is non-negative, $L_t = \sup_{\pi\in C} L_t^\pi$ implies $-\eta L_t = \inf_{\pi\in C}\{-\eta L_t^\pi\}$. Finally, using the fact that the function exp is bijective and increasing, we have the relations $-\exp\{-\eta L_t\} = -\inf_{\pi\in C}\exp\{-\eta L_t^\pi\} = \sup_{\pi\in C}(-\exp\{-\eta L_t^\pi\})$. Thus,

$$V^F(t,x) = -\exp\{-\eta L_t\}.$$

In particular, $V^F(x) = V^F(0,x) = -\exp\{-\eta L_0\}$. □

Remark 2.3.17. *The method discribed above enables us to find a closed form formula for the value function of the control problem in terms of the solution of a BSDE as in the case of the method using the martingale optimality principle. Note that the BSDEs characterizing the problem are different, but all with quadratic growth in z (at least up to a transformation). However, the latter method does not describe the optimal control of the problem.*

2.4 Discussion

In this chapter we have presented three methods to represent the solutions of our control problem by a BSDE. In Subsection 2.3.1 we applied the martingale optimality principle to the case of exponential utility, and the case of power and logarithmic utility. The differences between the two cases are that we did not include the consumption process in the case of exponential utility and we did not include the terminal liability in the case of CRRA utilities. Considering the consumption process in the case of exponential utility will lead to pretty much the same results by the same reasoning. The issue of terminal liability is trickier. If F is any non-zero $\mathcal{F}_T$-random variable, the cases of CRRA utilities become difficult to handle. Of course, for the power utility, the transformation made to write the utility from the terminal wealth as the product of a constant and an exponential is no longer possible for F non-zero. With respect to the logarithmic utility, it will be difficult to show that the process defined by

$$K_t^{\pi,c} = h(t)(\log(X_t^{\pi,c} - F) - Y_t) + \int_0^t \alpha \log(\tilde{c}_s)\, ds$$

is a supermartingale because there will be no cancellations of $X_t^{\pi,c} - F$ when using Itô's formula. Even Doob-Meyer multiplicative decomposition seems hard to apply here. Nevertheless, it is still possible to carry out the same sort of reasoning with a certain class of non-zero terminal liability as explained by Imkeller *et al.* (2011). They consider liabilities of the form

$$F = c + \int_0^T \xi_s \frac{dS_s}{S_s},$$

where c is a constant and $(\xi_t)_{t\in[0,T]}$ an adapted and square integrable process. Therefore, the terminal wealth becomes

$$X_T^{\pi,c} - F = x_F + \int_0^T \pi_s^F \frac{dS_s}{S_s} - \int_0^T c_s\, ds,$$

with $x_F = x - c$ and $\pi_t^F = \pi_t - \xi$, $0 \le t \le T$, in such a way that $X_T^{\pi,c} - F$ can be manipulated as if F was zero (with a new "investment strategy"). This can also be applied to our case of closed set of constraints $C \subset L^2(0,T;L(\Omega))$. In fact, if we endow $L^2(0,T;L(\Omega))$ with its topology of vector space, the new constraint set $C - \xi$ is still a closed set, a property that was of capital importance in the analysis of Subsection 2.3.1. This class of terminal liabilities has the further advantage that the described methods are still valid even if F is unbounded, see Imkeller *et al.* (2011). In general, if the terminal liability is not bounded, the uniqueness of the solutions of the characterizing BSDE is not guaranteed, see Theorems 3.3.7 and 3.3.9. Another feature of each of the methods is the information it gives about the optimal policies. The existence is out of question here, it is given by the asymptotic elasticity condition on the utility function. The method of Subsection 2.3.3 does not describe the optimal policies, whilst the characterisation by martingale optimality gives a non-empty set in which the optimal strategies lie. For the case where the consumption c is zero, the characterisation via Itô-Ventzell's formula gives a closed form formula for the optimal investment strategy. It is worth noting that the characterization of Subsection 2.3.3 is designed to work for the exponential utility, whilst the two other methods can be used with CARA and CRRA utility functions. The characterization via Itô-Ventzell's formula is even more general, as Equation (2.3.23) is valid for any utility functions with some nice properties (see Mania and Tevzadze (2008) for more details).

There is an approach to solving of stochastic control problems leading to BSDEs that we have not presented in our study: the stochastic maximum principle. It was introduced by Pontryagin in the 1960s for deterministic systems. Pontryagin's maximum principle was extended to stochastic systems by Bismut, Bensoussan, Kushner in the 1970s. We refer to Yong and Zhou (1999), Chapter 3 for more details about the stochastic maximum principle in stochastic control.

The next chapter is dedicated to the study of some properties of quadratic BSDEs driven by Brownian motion.

Chapter 3

Quadratic BSDEs Driven by Brownian Motion

Depending on the form (with respect to Z) of the integrand of the Lebesgue integral in the BSDE, the equation is said to be linear, Lipschitz or quadratic in Z (for the most commonly studied cases). In financial applications, especially in the Black-Scholes framework, the BSDE is usually driven by Brownian motion. However, more general drivers like martingales or Lévy processes can be considered. As seen in the previous chapter, quadratic BSDEs driven by Brownian motion play an important role in utility maximization problems.

3.1 Introduction

In this section, we define some of the concepts which will be used most often in the rest of the thesis. Throughout the chapter, $(\Omega, \mathcal{F}, P)$ is a probability space carrying a d-dimensional Brownian motion $(B_t)_{t\in[0.T]}$, $\mathbb{F}$ the augmented filtration. We call n-dimensional backward stochastic differential equation (BSDE) an equation of the form

$$dY_t = -f(t, Y_t, Z_t)\,dt + Z_t\,dB_t, \quad Y_T = \xi. \tag{3.1.1}$$

Where the terminal value of the process $(Y_t)_{t\in[0,T]}$ is a $\mathcal{F}_T$-random variable, and the function $f : \Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^{n\times d} \to \mathbb{R}^n$ is generally called the *generator* (or *driver*) of the BSDE, and is an integrable $\mathcal{F}\otimes\mathcal{B}^n\otimes\mathcal{B}^{n\times d}$-measurable function. The random variable ξ together with the function f are called parameters of the BSDE.

Before going any further in this introductory section, let us introduce notation for some spaces and norms.

3.1.0.1 Spaces

Let $p > 1$.

- $L^p(\mathbb{R}^n)$, the space of random variables $X : \Omega \mapsto \mathbb{R}^n$ normed by $\|X\|_{L^p} = E[|X|^p]^{\frac{1}{p}}$, and $L^\infty(\mathbb{R}^n)$, the space of essentially bounded random variables normed by $\|X\|_{L^\infty} = \operatorname{ess\,sup}_{\omega\in\Omega} |X(\omega)|$

- $\mathcal{S}^p(\mathbb{R}^n)$, the space of all predictable processes $(Y_t)_{t\in[0,T]}$ with values in $\mathbb{R}^n$ normed by $\|Y\|_{\mathcal{S}^p} = E\left[\left(\sup_{t\in[0,T]} |Y_t|\right)^p\right]^{\frac{1}{p}}$, and $\mathcal{S}^\infty(\mathbb{R}^n)$, the space of essentially bounded predictable processes

- BMO, the class of BMO-martingales normed by $\|.\|_{BMO}$ as defined in Appendix A,

- $H^p(\mathbb{R}^n)$, the space of all predictable processes $(X_t)_{t\in[0,T]}$ normed by $\|X\|_{H^p(\mathbb{R}^n)} = E\left[\left(\int_0^T |X_t|^2\,dt\right)^{\frac{p}{2}}\right]^{\frac{1}{p}}$

- $H^2_\beta(\mathbb{R}^n)$, the space of all predictable processes normed, for a given $\beta \in \mathbb{R}^+$, by $\|\phi\|_\beta = E\left[\int_0^T e^{\beta t}|\phi_t|^2\,dt\right]^{\frac{1}{2}}$

- $H^2_{n,d}(\mathbb{R}^n)$, the space of $n \times d$-dimensional predictable processes normed by[1] $\|Z\|^2_{H^2_{n,d}(\mathbb{R}^n)} = E\left[\int_0^T \langle Z_t, Z_t'\rangle\,dt\right]$

- E, the class of real-valued processes Y which are RCLL and such that $Y^* = \sup_{t\in[0,T]} |Y_t|$ has exponential moments of all orders, i.e. $\forall\lambda > 0$, $E\left[e^{\lambda Y^*}\right] < +\infty$

- $\mathbb{D}^{k,p}(\mathbb{R}^n)$, the set of Malliavin differentiable random variables normed, for a given $k \in \mathbb{N}$, by $\|\xi\|^p_{k,p} = \|\xi\|^p_{L^p} + \sum_{i=1}^k \||D^{(i)}\xi|\|^p_{(H^p)^i}$. With the operator $D^{(i)}$ defined in Subsection 3.4.

We shall call solution of (3.1.1) an adapted process $(Y, Z) = (Y_t, Z_t)_{t\in[0,T]}$ such that Y is a $\mathbb{R}^n$-valued continuous and adapted process, Z is a $\mathbb{R}^{n\times d}$-valued predictable process such that $M^Z = \int_0^\cdot Z_s\,dB_s$ is a BMO-martingale and the function $t \mapsto f(t, Y_t, Z_t)$ is integrable.

Definition 3.1.1. *A function f is said to grow quadratically, or to be quadratically non-linear if*

$$|f(\omega, t, y, z)| \leq a + b|z| + c|z|^2,$$

where a, b and c are positive constants. The BSDE (3.1.1) *will be said to be quadratic, or with quadratic growth if the function f grows quadratically.*

The BSDE (3.1.1) describes the stochastic dynamics of the process Y controlled by Z. Therefore, Z is often referred to as the control variable and Y as the value process. Thus, a BSDE's generator with the growth of Definition

[1] $\langle u, v\rangle = trace(uv')$, where u and v are vectors of compatible dimensions.

3.1.1 is said to be quadratically non-linear in the control variable. Note that the case $f \equiv 0$ leads to $Y_t = \xi - \int_0^t Z_s\,dB_s$. The theory of BSDEs is actually a generalization of the well known martingale representation theorem. To exhibit this, let us recall the theorem.

Theorem 3.1.2 (Martingale representation theorem)**.** *If ξ is a real valued square integrable $\mathcal{F}_T$-random variable, then Y with $Y_t = E[\xi|\mathcal{F}_t]$ is in $H^2(\mathbb{R})$ and can be represented as a stochastic integral with respect to B of the (unique) predictable process $(Z_t)_{t\in[0,T]}$, with $E\left[\int_0^T |Z_s|^2\,ds\right] < \infty$. For $t \in [0,T]$, we have*

$$\begin{aligned} Y_t = E\left[\xi|\mathcal{F}_t\right] &= E[\xi] + \int_0^t Z_s\,dB_s, \quad \textit{(forward representation)} \\ &= \xi - \int_t^T Z_s\,dB_s, \quad \textit{(backward representation)}. \end{aligned}$$

Proof. See Karatzas and Shreve (1988), Theorem 3.4.15, page 182. □

Thus, $Y_t = \xi - \int_t^T Z_s\,dB_s$, $t \in [0,T]$ is a BSDE with driver $f \equiv 0$. Its solution is given by Theorem 3.1.2. Note that since Y has to be adapted, $Y_t = \xi$, $Z_t = 0$ cannot be taken as solution (unless if ξ is constant).

After having presented the rather simple case of generator-less BSDEs, let us briefly discuss the case of linear BSDEs.

3.1.1 Linear BSDEs

The BSDE introduced by Bismut (1973) was of the linear type, i.e. such that the generator f is linear. These equations were then used to describe the adjoint process with the stochastic maximum principle. It has been shown, see for instance El Karoui *et al.* (1997) or Quenez (1993), that this class of BSDEs is useful in the resolution of the problem of pricing of contingent claims in a complete market. In fact, the Black-Scholes-Merton problem can be considered in terms of linear BSDEs. We present here an example of pricing in a complete market by means of a BSDE.

Let ξ be a European contingent claim due at time T. A trader wants to hedge this contract, by investing according to a self-financing strategy $\pi' = (\pi^0, \pi^1, \ldots, \pi^n)$ in $n+1$ assets including a risk-less bond with instantaneous yield r_t and n risky assets with price dynamics

$$dP_t^i = P_t^i \left[b_t^i\,dt + \sum_{j=1}^n \sigma_t^{i,j}\,dB_t^i \right], \quad P_0^i = p^i.$$

We also suppose that the trader consumes at a given rate $(c_t)_{t\in[0,T]}$, which is assumed to be adapted and non-negative. The wealth process $V = \sum_{i=0}^n \pi^i$ satisfies

$$dV_t = (r_t V_t + \pi_t' \sigma_t \theta_t - c_t)\,dt + \pi_t' \sigma_t\,dB_t,$$

where the market price of risk, θ, given by the formula $\sigma_t\theta_t = (b_t - r_t\mathbf{1})$ for all t is a predictable and bounded-valued vector process, and $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^n$. For more details about the derivation, see Karatzas and Shreve (1988), Section 5.8 page 371. The drift vector b_t and the volatility matrix σ_t are predictable, positive and bounded processes such that σ is invertible with bounded inverse σ^{-1}. The claim ξ is attained if $V_T = \xi$. Thus, the hedging of ξ leads us to the linear BSDE

$$dV_t = f(t, V_t, Z_t)\, dt - Z_t\, dB_t, \quad V_T = \xi. \tag{3.1.2}$$

With $-Z_t = \pi_t'\sigma_t$ and $f(t, V_t, Z_t) = r_tV_t + Z_t\theta_t - c_t$. Theorem 3.2.3 ensures the existence and the uniqueness of solution of (3.1.2), and, in addition, specifies the integrability properties of the solution.

Following El Karoui *et al.* (1997), Section 1.2, there exists a process $(H_t)_{t\in[0,T]}$ which is a solution of

$$dH_t = -H_t\left[r_t\, dt + \theta_t'\, dB_t\right], \quad H_0 = 1,$$

such that the process $E\left[H_T\xi|\mathcal{F}_t\right]$ is a uniformly integrable non-negative martingale. In fact, applying Itô's formula to the function $g(t, x) = \log(x)$, we have

$$\begin{aligned} d\log(H_t) &= \frac{dH_t}{H_t} - \frac{1}{2H_t^2}H_t^2(\theta_s')^2\, dt \\ &= -(r_t\, dt + \theta_t'\, dB_t) - \frac{1}{2}\theta_t'^2\, dt. \end{aligned}$$

Thus H is given by

$$H_t = \exp\left\{-\int_0^t r_s\, ds - \int_0^t \theta_s' dB_s - \frac{1}{2}\int_0^t |\theta_s|^2\, ds\right\}.$$

Now we define the process V by

$$H_tV_t = M_t - \int_0^t c_sH_s\, ds$$

with $M_t = E\left[H_T\xi + \int_0^T c_sH_s\, ds|\mathcal{F}_t\right]$ which admits $P \otimes \lambda$ a.s., according to the martingale representation theorem, the unique representation $M_t = M_0 + \int_0^t R_s'\, dB_s$. Where $(R_t')_{t\in[0,T]}$ is a square integrable process. Note that

$$H_tV_t = E\left[H_T\xi + \int_t^T c_sH_s\, ds|\mathcal{F}_t\right]. \tag{3.1.3}$$

Hence, we have $H_TV_T = H_T\xi$ and using the representation of M,

$$\begin{aligned} H_TV_T &= M_T - \int_0^T c_sH_s\, ds \\ &= M_0 + \int_0^T R_s'\, dB_s - \int_0^T c_sH_s\, ds. \end{aligned}$$

This implies

$$\begin{aligned}H_tV_t &= H_T\xi - H_TV_T + H_tV_t\\ &= H_T\xi - M_0 - \int_0^T R'_s\,dB_s + \int_0^T c_sH_s\,ds + M_0 + \int_0^t R'_s\,dB_s - \int_0^t c_sH_s\,ds\\ &= H_T\xi + \int_t^T c_sH_s\,ds - \int_t^T R'_s\,dB_s.\end{aligned}$$

Let us define $\pi'_t\sigma_t = (H_t^{-1}R'_t + V_t\theta'_t)$. By Itô product formula,

$$d(H_tV_t) = H_t\,dV_t + V_t\,dH_t + d[H_t, V_t].$$

Because $d(H_tV_t) = R'_t\,dB_t - c_tH_t\,dt$, we have

$$H_t(\pi'_t\sigma_t - V_t\theta'_t)\,dB_t - c_tH_t\,dt = H_t\,dV_t - H_tV_t(r_t\,dt + \theta'_t dB_t) - H_t\theta'_t\,d[V_t, B_t].$$

This implies

$$\begin{aligned}dV_t &= (\pi'_t\sigma_t - V_t\theta'_t)\,dB_t + V_t(r_t\,dt + \theta'_t\,dB_t) + \theta'_t\,d[V_t, B_t] - c_t\,dt\\ &= \pi'_t\sigma_t\,dB_t + r_tV_t\,dt - c_t\,dt + \pi'_t\sigma_t\,dt.\end{aligned}$$

Therefore, $(V_t, \pi'_t\sigma_t)_{t\in[0,T]}$ solves the linear BSDE (3.1.2). The price of the contract is given by V_0.

The issue of well-posedness of linear BSDEs was first solved by Bensoussan (1983), using the martingale representation theorem and a contraction mapping argument. As a stochastic version of the Bellman equation, Chitashvili (1983) derived a non-linear semimartingale BSDE with a particular generator and established the well-posedness of BSDEs with Lipschitz generators. His work was extended, in a Brownian setting, by Pardoux and Peng (1990) to general Lipschitz generators. The primary goal when introducing the theory of BSDEs of non-linear form was, for Pardoux and Peng, to find a Feynman-Kac representation of solutions for a class of non-linear second-order PDEs. Thus, generator-less BSDEs generalise the martingale representation theorem whilst BSDEs (with non-zero generator) generalise the Feynman-Kac formula.

3.1.2 Feynman-Kac Formula

In this subsection, we present the link between solutions of BSDEs and solutions of a class of PDEs. This link is made by the so called Feynman-Kac formula, which gives a probabilistic interpretation to the solution of a PDE. We start by introducing the concept of forward backward stochastic differential equations (FBSDEs for short).

The terminal condition ξ can be given in terms of the solution of a stochastic differential equation, say $\xi = g(X_T)$ with g a given measurable function and $X^{u,x}$ the solution (starting at time u with value x) of the SDE

$$X_t^{u,x} = x + \int_u^t b(s, X_s)\,ds + \int_u^t \sigma(s, X_s)\,dB_s. \tag{3.1.4}$$

With $u \in [0,T]$, $t \in [u,T]$, $b : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times d}$ being two deterministic continuous functions satisfying the following assumptions:

(**SDE1**) there exists a constant $\beta \geq 0$ such that for all $(t,x,x') \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^n$, $b(t,0) \leq \beta$, $\sigma_i(t,0) \leq \beta$, $1 \leq i \leq d$ and

$$|b(t,x) - b(t,x')| + |\sigma(t,x) - \sigma(t,x')| \leq \beta|x - x'|$$

(**SDE2**) σ is bounded, the functions $x \mapsto b(.,x)$ and $x \mapsto \sigma_i(.,x)$, $(1 \leq i \leq d)$ are continuously differentiable and their derivatives satisfy the standard Lipschitz condition in x with Lipschitz constant β.

We consider the BSDE

$$Y_t = g(X_T^{u,x}) + \int_t^T f(s, X_s^{u,x}, Y_s, Z_s)\,ds - \int_t^T Z_s\,dB_s, \quad t \in [u,T]. \tag{3.1.5}$$

The system (3.1.5)-(3.1.4) is a FBSDE. Even though we still write Y_t and Z_t, one should notice that the processes Y and Z actually depend on the initial state x of the process X, and the initial time u. Classical results on SDEs show that under assumptions (SDE1) and (SDE2), Equation (3.1.4) has a unique solution $X \in \mathcal{S}^p(\mathbb{R}^n)$ for all $p \geq 1$. We assume that Equation (3.1.5) has a solution. Therefore the system (3.1.5)-(3.1.4) has at least one solution denoted by (X,Y,Z). In the case of linear BSDEs, more precisely, if we consider the BSDE

$$Y_t = g(X_T^{u,x}) + \int_t^T \big(b(s, X_s^{u,x})Y_s + \sigma(s, X_s^{u,x})Z_s\big)\,ds - \int_t^T Z_s\,dB_s, \tag{3.1.6}$$

with $t \in [u,T]$. Using an extension of the classical variation of the constant, Pardoux shows that (see the proof of Theorem 1.6. in Pardoux (1996)) the process Y in the solution of (3.1.6) is given by

$$\begin{aligned} Y_t = g(X_t^{u,x}) \exp\left(\int_t^T b(s, X_s^{u,x})ds\right) &+ \int_t^T \sigma(s, X_s^{u,x}) \exp\left(\int_t^s b(r, X_r^{u,x})\,dr\right) ds \\ &- \int_t^T \exp\left(\int_t^s b(r, X_r^{u,x})\,dr\right) Z_s\,dB_s. \end{aligned}$$

Taking the conditional expectation of both sides and using the fact that Y is adapted yields

$$\begin{aligned} Y_t = E_{t,x}&\left[g(X_t^{u,x}) \exp\left(\int_t^T b(s, X_s^{u,x})\,ds\right)\right. \\ &\left.+ \int_t^T \sigma(s, X_s^{u,x}) \exp\left(\int_t^s b(r, X_r^{u,x})\,dr\right)\,ds\right] \\ &= v(t,x), \end{aligned}$$

where the second equality is given by the Feynman-Kac representation of the unique solution of the second order PDE

$$-\frac{\partial v}{\partial t} + bv = \frac{1}{2}\Delta v + \sigma \quad \text{in} \quad [0, T) \times \mathbb{R}^n,$$

with boundary condition $v(T, x) = g(x)$.

The above representation of the solution of a BSDE in terms of the solution of a PDE can be extended, as shown for instance by Kobylanski (2000), to the more general case of non-linear BSDEs.

3.2 Digression into the Lipschitz Continuous Case

In this section we open a parenthesis to give an overview of the principal results on BSDEs with Lipschitz drivers. The section is mainly inspired by El Karoui *et al.* (1997) where the authors highlight the applications of BSDEs in finance, notably the theory of recursive utilities and pricing of contingent claims.

Remark 3.2.1. *In this section the parameters of the BSDE are such that* $\xi \in L^2(\mathbb{R}^n)$*,* $f(.,0,0) \in H^2(\mathbb{R}^n)$*, and* f *is uniformly Lipschitz; i.e. there exists a constant* $C > 0$ *such that* $\lambda \otimes P$ *a.s., we have*

$$|f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)| \leq C(|y_1 - y_2| + |z_1 - z_2|) \quad \forall y_1, y_2, z_1, z_2.$$

The following proposition provides a key tool in the proof of the existence and uniqueness of solution of a BSDE with a Lipschitz driver.

Proposition 3.2.2. *Let* (f^1, ξ^1) *and* (f^2, ξ^2) *be two pairs of parameters of the BSDE. Assume the existence of* (Y^1, Z^1) *and* (Y^2, Z^2) *two square-integrable solutions. Let* f^1 *be a Lipschitz function with Lipschitz constant* C*, and put* $\delta Y_t = Y_t^1 - Y_t^2$ *and* $\delta_2 f_t = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)$*. For any* (λ, μ, β) *such that* $\mu > 0$*,* $\lambda^2 > C$ *and* $\beta \geq C(2 + \lambda^2) + \mu^2$*, it follows that*

$$\begin{aligned}
\|\delta Y\|_\beta^2 &\leq T \left[e^{\beta T} E[|\delta Y_T|^2] + \frac{1}{\mu^2} \|\delta_2 f\|_\beta^2 \right] \\
\|\delta Z\|_\beta^2 &\leq \frac{\lambda^2}{\lambda^2 - C} \left[e^{\beta T} E[|\delta Y_T|^2] + \frac{1}{\mu^2} \|\delta_2 f\|_\beta^2 \right].
\end{aligned}$$

Proof. See El Karoui *et al.* (1997), Proposition 2.1. □

Theorem 3.2.3 (Pardoux-Peng)**.** *Given two parameters such as in Remark 3.2.1, there exists a unique pair of processes* $(Y, Z) \in H^2(\mathbb{R}^n) \times H^2(\mathbb{R}^{n \times d})$ *which solves BSDE* (3.1.1)*.*

Sketch of the Proof. The reader may find a detailed proof in El Karoui *et al.* (1997), Theorem 2.1. We explain here the idea of the proof.

Define a map $\Phi : H^2_\beta(\mathbb{R}^n) \times H^2_\beta(\mathbb{R}^n) \mapsto H^2_\beta(\mathbb{R}^n) \times H^2_\beta(\mathbb{R}^n)$ which maps a point (y, z) onto the process (Y, Z) such that

$$Y_t = \xi + \int_t^T f(s, y_s, z_s)\, ds - \int_t^T Z_s\, dB_s.$$

The map Φ is well defined, since the assumptions on f in Remark 3.2.1 imply that the process $(f(t, y_t, z_t))_{t\in[0,T]}$ is integrable with respect to t. Moreover, the same remark implies that $(f(t, y_t, z_t))_{t\in[0,T]}$ and ξ are square-integrable. So, we can use the martingale representation theorem to justify the existence of a unique square integrable process Z such that the (continuous version of the) square-integrable martingale $M_t = E\left[\int_0^T f(s, y_s, z_s)\, ds + \xi|\mathcal{F}_t\right]$ is written as $M_0 + \int_0^t Z_s\, dB_s$. Putting

$$Y_t = M_t - \int_0^t f(s, y_s, z_s) dB_s = E\left[\xi + \int_t^T f(s, y_s, z_s)\, dB_s \mid \mathcal{F}_t\right],$$

one easily derives $Y_t = \xi + \int_t^T f(s, y_s, z_s)\, ds - \int_t^T Z_s\, dB_s$. Which shows that the pair (Y, Z) is uniquely defined. Now, using the above a priori estimate with suitable parameters, and the fact that f is Lipschitz, one shows that Φ is a contraction of the Banach space $H^2_\beta(\mathbb{R}^n) \times H^2_\beta(\mathbb{R}^n)$ onto itself. The fixed point theorem gives existence and uniqueness of a solution. □

One can see that the assumption that f is Lipschitz with constant coefficient is a key argument in the proof. In addition, the terminal condition ξ needs not to be bounded, but only square-integrable.

We give another result.

Proposition 3.2.4 (Comparison)**.** *Let f^1, ξ^1 and f^2, ξ^2 be parameters of two given BSDEs, respectively. Let (Y^1, Z^1) and (Y^2, Z^2) be the respective square-integrable solutions. Assume that the inequalities*[2] *$\xi^1 \geq \xi^2$ a.s., and $f^1(t, Y^2_t, Z^2_t) \geq f^2(t, Y^2_t, Z^2_t)$ $\lambda \otimes P$ a.s. hold. Then, for any time t, we have $Y^1_t \geq Y^2_t$. If in addition $Y^1_t = Y^2_t$ on a set $A \in \mathcal{F}_t$, then $Y^1_s = Y^2_s$ a.s. on $[t, T] \times A$, $\xi_1 = \xi_2$ a.s. on A, and $f^1(t, Y^2_t, Z^2_t) = f^2(t, Y^2_t, Z^2_t)$ on $A \times [t, T]$ $\lambda \otimes P$ a.s.*

Proof. See El Karoui *et al.* (1997), Theorem 2.2. □

Remark 3.2.5. *An important financial consequence of this proposition is given by the next corollary. It shows that the price of the claim ξ in the study of Subsection 3.1.1 does not lead to an arbitrage. In fact, if the claim and the consumption rate satisfy $\xi \geq 0$ and $c_t \geq 0$ $\lambda \otimes P$ a.s., then the price V_0 should be non-negative. Conversely, if $V_0 = 0$, then the value of the claim should also be $\xi = 0$.*

[2]The inequality $x \geq y$ for two vectors of same dimension should be understood as $x^i \geq y^i$ for all i, i.e. componentwise.

Corollary 3.2.6. *If $\xi \geq 0$ a.s. and $f(t,0,0) \geq 0$ $\lambda \otimes P$ a.s., then $Y \geq 0$ a.s. In addition, if $Y_t = 0$ on a set $A \in \mathcal{F}_t$, then $Y_s = 0$, $f(s,0,0) = 0$ on $[t,T] \times A$, $\lambda \otimes P$ a.s., and $\xi = 0$ a.s. on A.*

Proof. See El Karoui *et al.* (1997). □

The following result describes the properties of differentiability of the solution of a BSDE with Lipschitz generator.

Let $(f_\alpha, \alpha \in \mathbb{R})$, and $(\xi_\alpha, \alpha \in \mathbb{R})$ be two families of parameters of BSDEs such that f_α is uniformly Lipschitz and $\xi_\alpha \in L^2(\mathbb{R}^n)$ for all α. Let (Y_α, Z_α) be the solution of the BSDE with parameters f_α and ξ_α. Suppose that there exists $C > 0$ such that $\lambda \otimes P$ a.s.,

$$\forall \alpha, \quad |f_\alpha(\omega,t,y_1,z_1) - f_\alpha(\omega,t,y_2,z_2)| \leq C(|y_1 - y_2| + |z_1 - z_2|),$$

and that for each α_0, $f_\alpha(t, Y_t^{\alpha_0}, Z_t^{\alpha_0}) - f_{\alpha_0}(t, Y_t^{\alpha_0}, Z_t^{\alpha_0}) \to 0$ in $H^2_\beta(\mathbb{R}^n)$ and $\xi_\alpha - \xi_{\alpha_0} \to 0$ in $L^2(\mathbb{R}^n)$.

Proposition 3.2.7. *Further assume that the families of parameters $(f_\alpha, \alpha \in \mathbb{R})$ and $(\xi_\alpha, \alpha \in \mathbb{R})$ are such that, for all $\alpha \in \mathbb{R}$, f_α is differentiable with respect to (y,z) and has a uniformly bounded and continuous partial derivative in x and y, and that the functions $\alpha \mapsto f_\alpha(., Y^\alpha, Z^\alpha)$ and $\alpha \mapsto \xi^\alpha$ are differentiable.*

Then, the function $\alpha \mapsto (Y^\alpha, Z^\alpha)$, from $\mathbb{R}$ to $H^2_\beta(\mathbb{R}^n) \times H^2_\beta(\mathbb{R}^n)$ is differentiable with derivative given by $(\partial_\alpha Y^\alpha, \partial_\alpha Z^\alpha)$, the solution of the BSDE

$$\begin{aligned} -d(\partial_\alpha Y_t^\alpha) &= \langle \partial f_\alpha(t, Y_t^\alpha, Z_t^\alpha), (1, \partial_\alpha Y_t^\alpha, \partial_\alpha Z_t^\alpha) \rangle \, dt - (\partial_\alpha Z^\alpha)' \, dB_t \\ \partial_\alpha Y_T^\alpha &= \partial_\alpha \xi^\alpha, \end{aligned}$$

with $\partial f_\alpha = (\partial_\alpha f_\alpha, \partial_y f_\alpha, \partial_z f_\alpha)'$.

Proof. See El Karoui *et al.* (1997), Proposition 2.4. □

3.3 Generalities on BSDEs with Quadratic Growth

This section is dedicated to the presentation of some results from the theory of BSDEs with quadratic drivers. The stochastic approach used in Chapter 2 to deal with a control problem such as (1.2.7) entailed a (one-dimensional) BSDE with a quadratic driver. Moreover, in some financial problems like the problem of equilibrium when we consider the utility maximization of many investors acting together, the setting leads to a multi-dimensional quadratic BSDE, see Frei and Reis (2011). These problems encountered in financial mathematics have stimulated a growing attention to this class of equations, and the need for a definitive answer to the question of existence and uniqueness of solutions in a general case. Kobylanski (2000) proved the first ever existence and uniqueness result for quadratic BSDEs. She considered a one-dimensional BSDEs with bounded terminal condition. However, the rapid evolution of financial applications has given rise to more diverse quadratic BSDEs, for which the hypotheses of Kobylanski are too restrictive. Before giving an overview of some results on quadratic BSDEs, let us give this counterexample, taken from Frei and Reis (2011), which gives a flavour of the challenge to find a general existence and uniqueness result.

3.3.1 Counterexample

Let us start by giving the following result which will be used in the argument below.

Lemma 3.3.1. *There exists* $\kappa \in H^2_{1,1}$, *with*

$$\int \kappa\, dB \in \mathcal{S}^\infty \quad \textit{and} \quad E\left[\exp\left(\int_0^T |\kappa_t|^2\, dt\right)\right] = \infty.$$

Proof. See Frei and Reis (2011), Appendix A.1. □

We consider the two-dimensional BSDE

$$dY_t = -f(t, Y_t, Z_t)\, dt + Z_t\, dW_t, \quad Y_T = \xi, \tag{3.3.1}$$

with $Y = (Y^1 \;\; Y^2)$, $Z = (Z^1 \;\; Z^2)$ and

$$f(t, Y_t, Z_t) = \begin{pmatrix} 0 \\ |Z_t^1|^2 + \frac{1}{2}|Z_t^2|^2 \end{pmatrix}, \quad \xi = \begin{pmatrix} \xi_1 \\ 0 \end{pmatrix}.$$

Moreover, $W_t = (B_t \;\; B_t)'$ for all $t \in [0, T]$, and $\xi_1 \in L^\infty(\mathbb{R})$ is given.

The driver f in Equation (3.3.1) obviously growths quadratically in the control variable. In fact, if we consider the Euclidean norm $|.|$, for all t we have

$$|f(t, Y_t, Z_t)| = |Z_t^1|^2 + \frac{1}{2}|Z_t^2|^2 \le |Z_t^1|^2 + |Z_t^2|^2 = |Z_t|^2.$$

By definition, $dY_t^1 = Z_t^1\, dB_t$ and $Y_T^1 = \xi_1$. This implies $Y_t^1 = E\left[\xi_1 | \mathcal{F}_t\right]$, and by Theorem 3.1.2, there exists a unique square-integrable process Z^1 such that $Y_t^1 = Y_0^1 + \int_0^t Z_s^1\, dB_s$. In addition, $Y_T^1 = E\left[\xi_1 | \mathcal{F}_T\right] = \xi_1$ and $Y_0^1 = E\left[\xi_1 | \mathcal{F}_0\right] = E[\xi_1]$, hence

$$\xi_1 = E[\xi_1] + \int_0^T Z_s^1\, dB_s. \tag{3.3.2}$$

Having Z^1, from $dY_t^2 = -(|Z_t^1|^2 + \frac{1}{2}|Z_t^2|^2)\, dt + Z_t^2\, dB_t$, we can write

$$\begin{aligned}\exp\left(\int_0^T |Z_t^1|\,dt\right) &= \exp(Y_0^2)\exp\left(-\frac{1}{2}\int_0^T |Z_t^2|^2\,dt + \int_0^T Z_t^2\,dB_t\right)\\ &= \exp(Y_0^2)\mathcal{E}\left(\int Z^2\,dB\right)_T. \qquad (3.3.3)\end{aligned}$$

Because the stochastic exponential $\mathcal{E}\left(\int Z^2\,dB\right)$ is a positive supermartingale, for all $t \geq 0$ we have

$$E\left[\mathcal{E}\left(\int Z^2\,dB\right)_t\right] \leq E\left[\mathcal{E}\left(\int Z^2\,dB\right)_0\right] = 1.$$

Hence, taking the expectation both sides in (3.3.3) leads to the inequality

$$E\left[\exp\left(\int_0^T |Z_t^1|\,dt\right)\right] \leq \exp(Y_0^2). \qquad (3.3.4)$$

Put $\xi_1 = \int_0^T \kappa_t\,dB_t \in L^\infty$, with κ defined by Lemma 3.3.1. Equation (3.3.2) implies $Z^1 = \kappa$ (use for instance Itô isometry). From Lemma 3.3.1 and Equation (3.3.4), we have $Y_0^2 = \infty$. Therefore, there does not exist a process (Y, Z) which solves BSDE (3.3.1) in the sence defined in the introductory section.

3.3.2 Existence, Uniqueness and Stability

These three issues: existence, uniqueness and stability of solutions of BSDEs are at the core of research in quadratic BSDEs. We devote this subsection to an overview of some results currently available in the literature.

3.3.2.1 Existence

The first theoretical results on BSDEs with quadratic growth in the control variable were published at the end of the nineties by Kobylanski , see Kobylanski (1997, 2000). Her existence result is the following.

Theorem 3.3.2 (Kobylanski)**.** *Assume that ξ is a $\mathcal{F}_T$-measurable random variable absolutely bounded by a real number $K > 0$ and the function $f : \Omega \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ is adapted, measurable and continuous in the spatial variable. Assume further that there exists a real constant $M > 0$ such that for any $(\omega, t) \in \times[0, T]$, $y, y' \in \mathbb{R}$ and $z, z' \in \mathbb{R}^d$ one has*

$$|f(t, y, z)| \leq M(1 + |y| + c(|y|)|z|^2) \quad a.s.,$$

with $c : \mathbb{R}^+ \to \mathbb{R}^+$ continuous and increasing.

Then BSDE (3.1.1) *has at least one solution $(Y, Z) \in \mathcal{S}^\infty(\mathbb{R}) \times H^2(\mathbb{R}^d)$ such that the process Y has continuous paths.*

Furthermore, there exists a minimal solution (Y_*, Z_*) such that for any BSDE with parameters g, ζ, if

$$f \leq g \quad \text{and} \quad \xi \leq \zeta,$$

for any solution (Y_g, Z_g) of the BSDE with parameters g and ζ, we have

$$Y_* \leq Y_g.$$

An analogous result holds for maximal solutions.

Proof. For the details of the proof, see Kobylanski (2000), Theorem 2.3. In brief, the proof of this theorem is heavily based on the monotone stability result of Proposition 3.3.14. To establish her proof, Kobylanski makes an exponential change of variable to transform the BSDE (3.1.1) into another one. Then she constructs an approximation of the generator of the new BSDE by a sequence of uniformly Lipschitz continuous functions and uses Proposition 3.3.14. □

In the particular case where the generator takes the form

$$f(c, y, z) = c - \beta y - \frac{\alpha}{2} z^2, \quad \alpha, \beta \geq 0.$$

With terminal condition $\xi = 0$, Schroder and Skiadas (1999) prove an existence and uniqueness result using a fixed point theorem. This class of BSDE has been shown to define an important form of stochastic differential utility.

Example 3.3.3. *We give here the example of the following well know BSDE that Kobylanski used to justify the importance of the boundedness assumption of ξ, and to explain the technique of exponential change of variable.*

Consider the equation

$$Y_t = \xi - \frac{1}{2} \int_t^T |Z_s|^2 \, ds - \int_t^T Z_s \, dB_s \quad 0 \leq t \leq T. \tag{3.3.5}$$

Put $P_t = e^{Y_t}$. Taking the exponential both sides in (3.3.5), we have

$$\begin{aligned} P_t &= e^{\xi} \exp\left\{ -\frac{1}{2} \int_t^T |Z_s|^2 \, ds - \int_t^T Z_s \, dB_s \right\} \\ &= e^{\xi} \mathcal{E}(M)_t, \end{aligned}$$

where M is the process given by $M_t = -\int_t^T Z_s \, dB_s$. By definition of the stochastic exponential, we have

$$P_t = e^{\xi} \left(1 + \int_t^T (-Z_s) \mathcal{E}(M)_s \, dB_s \right).$$

Thus, BSDE (3.3.5) *becomes*

$$P_t = e^{\xi} - \int_t^T Q_s \, dB_s, \tag{3.3.6}$$

with $Q_t = e^{Y_t} Z_t$, $t \in [0,T]$. *Since Equation* (3.3.6) *is linear, we have from Theorem 3.2.3 that if* $e^{\xi} \in L^2(\mathbb{R})$, *BSDE* (3.3.6) *admits a unique solution* $(P,Q) \in H^2(\mathbb{R}) \times H^2(\mathbb{R}^d)$, *with* $P_t = E\left[e^{\xi}|\mathcal{F}_t\right]$, $t \in [0,T]$, *hence*

$$Y_t = \log(E\left[e^{\xi}|\mathcal{F}_t\right]). \tag{3.3.7}$$

Kobylanski pointed out that in order to have $e^{\xi} \in L^2(\mathbb{R})$, *a sufficient condition is* $\xi \in L^{\infty}(\mathbb{R})$.

The condition $\xi \in L^{\infty}(\mathbb{R})$ is too strong, and appears less naturally than $E[e^{\xi}] < \infty$, which in turn is less restrictive. Based on this observation, Briand and Hu (2006) show that for this case, the existence of an exponential moment of ξ is sufficient to construct a solution. Let us define by (Y^n, Z^n), the minimal solution of the following quadratic BSDE:

$$Y_t^n = \xi \wedge n - \frac{1}{2}\int_t^T |Z_s^n|^2 \, ds - \int_t^T Z_s^n \, dB_s \quad t \in [0,T],$$

with bounded terminal condition $\xi \wedge n$. From (3.3.7), we have

$$-\log(E\left[e^{-(\xi\wedge n)}|\mathcal{F}_t\right]) \le Y_t^n \le \log(E\left[e^{(\xi\wedge n)}|\mathcal{F}_t\right]), \quad \text{for all} \quad n,t.$$

One can prove the first inequality by contrapositive and appeal to Jensen's inequality. Let $\tau_k = \inf\left\{t \in [0,T] : \log(E\left[e^{\xi}|\mathcal{F}_t\right]) \ge k\right\} \wedge T$, $k \in \mathbb{N}$, and consider the BSDE

$$Y_{t\wedge\tau_k}^n = Y_{\tau_k}^n - \frac{1}{2}\int_{t\wedge\tau_k}^{T\wedge\tau_k} |Z_s^n|^2 \, ds - \int_{t\wedge\tau_k}^{T\wedge\tau_k} Z_s^n \, dB_s, \quad t \in [0,T]$$

which is equivalent to restricting the study to the random interval $[0,\tau_k]$. By definition of τ_k, we have $\|Y_{t\wedge\tau_k}^n\|_{\infty} \le k$ for all n and t. In addition $(Y^n)_{n\in\mathbb{N}}$ is increasing by Theorem 3.3.2. Hence for a fixed k the sequence $\{(Y_{t\wedge\tau_k}^n)_{t\in[0,T]} : n \in \mathbb{N}\}$ converges to its upper bond (see Proposition 3.3.14) then we obtain the solution by sending k to infinity. We recall that ξ was not assumed to be bounded to carry out the above analysis, but we only assumed that $E\left[e^{\xi}\right] < \infty$. Briand and Hu generalise this reasoning to prove the following result:

Theorem 3.3.4 (Briand-Hu). *Assume that* f *is continuous with respect to* (y,z), *and that for all* $(t,y,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}^d$,

$$|f(t,y,z)| \le \alpha + \beta|y| + \frac{\gamma}{2}|z|^2, \quad \alpha,\beta \ge 0, \ \gamma > 0;$$

and that there exists $\lambda > \gamma$ such that

$$E\left[e^{\lambda e^{\beta T}|\xi|}\right] < \infty. \tag{3.3.8}$$

Then, BSDE (3.1.1) *has at least one solution $(Y,Z) \in \mathcal{S}^\infty(\mathbb{R}) \times H^2(\mathbb{R}^d)$ such that*

$$-\frac{1}{\gamma}\log\left(E\left[\phi_t(-\xi)|\mathcal{F}_t\right]\right) \leq Y_t \leq \frac{1}{\gamma}\log\left(E\left[\phi_t(\xi)|\mathcal{F}_t\right]\right),$$

where $(\phi_t(z))_{t\in[0,T]}$ denotes the solution of the equation

$$\phi_t = e^{\gamma z} + \int_t^T H(\phi_s)\,ds \quad t \in [0,T],$$

and $H(p) = p(\alpha\gamma + \beta\log(p))1_{[1,+\infty)}(p) + \gamma\alpha 1_{(-\infty,1)}(p)$.

Proof. See Briand and Hu (2006), Section 4. □

The proof of Theorem 3.3.4 is based on the following estimate for Y.

Lemma 3.3.5. *If f is as in the assumption of Theorem 3.3.4, ξ bounded, then the solution of the BSDE satisfies*

$$-\frac{1}{\gamma}\log\left(E\left[\phi_t(-\xi)|\mathcal{F}_t\right]\right) \leq Y_t \leq \frac{1}{\gamma}\log\left(E\left[\phi_t(\xi)|\mathcal{F}_t\right]\right).$$

Proof. See Briand and Hu (2006), Lemma 1. □

We now give the following example, inspired from Ankirchner *et al.* (2009), which shows that the integrability condition (3.3.8) is the best possible condition for existence of solutions.

Example 3.3.6. *Consider the BSDE*

$$Y_t = \frac{B_1^2}{2} + \int_t^1 \frac{1}{2}Z_s^2\,ds - \int_t^1 Z_s\,dB_s, \quad t \in [0,1], \tag{3.3.9}$$

with parameters $\xi = \frac{B_1^2}{2}$ and $f(t,Z_t) = \frac{1}{2}Z_t^2$. Since B_1 possesses a standard normal density, $E\left[\exp\left(\frac{B_1^2}{2}\right)\right] = \infty$. Hence, $E\left[\exp(\lambda|\xi|)\right] = \infty$ for all $\lambda > 1$.

Let

$$Z_t = \frac{B_t}{t} \quad \text{for} \quad t > 0 \quad \text{and} \quad Z_0 = 0.$$

By the Itô product formula, we have

$$Z_t dB_t = d(Z_t B_t) - B_t dZ_t - dZ_t dB_t.$$

For $t > 0$, by integration we have

$$\begin{aligned}\int_t^1 Z_s dB_s &= B_1^2 - \frac{B_t^2}{t} - \int_t^1 B_s\left(-\frac{B_s}{s^2}\,ds + \frac{1}{s}\,dB_s\right)\\ &\quad - \int_t^1 \left\{-\frac{B_s}{s}\,d\left[s, B_s\right] + \frac{1}{s}\,d\left[B,B\right]_s\right\}\\ &= B_1^2 - \frac{B_t^2}{t} + \log(t) + \int_t^1 \frac{B_s^2}{s^2}\,ds - \int_t^1 \frac{B_s}{s}\,dB_s\\ &= \frac{1}{2}B_1^2 - \frac{1}{2}\left(\frac{B_t^2}{t} - \log(t)\right) + \int_t^1 \frac{B_s^2}{2s^2}\,ds.\end{aligned}$$

Thus, for all $t > 0$, the process $(Y,Z) = \left(\frac{1}{2}\left(\frac{B_s^2}{s} - \log(s)\right), \frac{B_s}{s}\right)_{s\in[t,1]}$, and $(Y_0, Z_0) = (0,0)$ solves BSDE (3.3.9) on $[t,1]$, $t > 0$. On the other hand, the process $\left(\frac{B_s}{s}\right)_{s\in[0,1]}$ is not square integrable on $[0,1]$, therefore (Y,Z) is not a solution of (3.3.9) on $[0,1]$. Consequently, Equation (3.3.9) does not admit solutions, because the local Lipschitz condition implies uniqueness on $[t,1]$ for all $t > 0$.

However, if we choose in BSDE (3.3.9) the terminal condition such that Condition (3.3.8) holds, solutions exist. That is what Ankirchner et al. *(2009) show by considering*

$$\xi = \frac{B_1^2}{2(1+\epsilon)}, \quad \epsilon > 0.$$

We have $\frac{1}{2(1+\epsilon)} < \frac{1}{2}$ and by the structure of the real line $\mathbb{R}$, there exists $\lambda > 1$ such that $\frac{\lambda}{2(1+\epsilon)} - \frac{1}{2} < 0$. Hence, $E\left[e^{\lambda|\xi|}\right] < \infty$. Put $Z_t = \frac{B_t}{t+\epsilon}$, $t \in [0,1]$. By the Itô product formula, we have

$$Z_t dB_t = d(Z_t B_t) - B_t dZ_t - dZ_t dB_t.$$

This implies for $t \in [0,1]$,

$$\begin{aligned}&\int_t^1 Z_s\,dB_s\\ &= \frac{B_1}{1+\epsilon} - \frac{B_t^2}{t+\epsilon} - \int_t^1 B_s\left(\frac{dB_s}{s+\epsilon} - \frac{B_s}{(s+\epsilon)^2}\,ds\right) - \int_t^1\left(\frac{dB_s}{s+\epsilon} - \frac{B_s}{s+\epsilon}\,ds\right)dB_s\\ &= \frac{B_1}{1+\epsilon} - \frac{B_t^2}{t+\epsilon} - \int_t^1 Z_s\,dB_s + \int_t^1 \frac{B_s^2}{(s+\epsilon)^2}\,ds - \int_t^1 \frac{ds}{s+\epsilon}\\ &= \frac{B_1}{2(1+\epsilon)} - \frac{1}{2}\left(\frac{B_t^2}{t+\epsilon} - \log\left(\frac{t+\epsilon}{1+\epsilon}\right)\right) + \int_t^1 \frac{1}{2}Z_s^2\,ds.\end{aligned}$$

Hence the process $\left(\frac{1}{2}\left(\frac{B_t^2}{t+\epsilon} - \log\left(\frac{t+\epsilon}{1+\epsilon}\right)\right), \frac{B_t}{t+\epsilon}\right)$ solves BSDE (3.3.9).

3.3.2.2 Uniqueness

The main argument of uniqueness of solution of BSDEs is the comparison results.

Theorem 3.3.7 (Comparison for BSDEs with bounded terminal condition)**.** *Let f^1, ξ^1 and f^2, ξ^2 be the parameters of two BSDEs. We suppose that*

- $\xi^1 \leq \xi^2$ *and* $f^1 \leq f^2$ *a.s.*
- *Either* f^1 *or* f^2 *satisfies*

$$|f(t,y,z)| \leq l(t) + c|z|^2 \quad a.s., \qquad \left|\frac{\partial f}{\partial z}(t,y,z)\right| \leq k(t) + c|z|^2 \quad a.s.,$$

and

$$\left|\frac{\partial f}{\partial y}(t,y,z)\right| \leq m(t) + c|z|^2 \quad a.s.$$

Where l, k, m *are functions of* t*, and* $c > 0$.

- ξ *is a* $\mathcal{F}_T$*-measurable random variable absolutely bounded.*

Then, if $(Y^i, Z^i) \in L^\infty(\mathbb{R}) \times H^2(\mathbb{R})$ *is a solution of the BSDE with parameters* f^i, ξ^i, $i = 1, 2$, *one has*

$$\forall t \in [0,T],\ Y_t^1 \leq Y_t^2 \ a.s.$$

Proof. See Kobylanski (2000), Theorem 2.6. □

The uniqueness of solution of the BSDE is a straightforward consequence of the above theorem.

Corollary 3.3.8. *Under the assumptions of Theorem 3.3.7, BSDE* (3.1.1) *has a unique solution* $(Y, Z) \in \mathcal{S}^\infty \times H^2(\mathbb{R}^d)$.

Proof. Let (Y', Z') and (Y, Z) be two solutions of (3.1.1) with parameters f and ξ like in Theorem 3.3.7. The theorem entails $Y' \leq Y$ and $Y \leq Y'$ a.s., i.e. $Y = Y'$ a.s. By a pathwise interpretation of the Itô's integral $\int_t^T (Z'_s - Z_s)\, dB_s$, we have $Z = Z'$ a.s., hence the uniqueness. □

In the case where ξ is not assumed to be bounded, Briand and Hu (2006) did not assess the uniqueness of solutions. They filled that gap two years later in Briand and Hu (2008), where they proved the comparison principle stated as follows.

Theorem 3.3.9 (Comparison for BSDEs with unbounded terminal condition)**.** *Let* (Y^1, Z^1) *(respectively* (Y^2, Z^2)*) be solution to the BSDE with parameters* f^1 *and* ξ^1 *(respectively* f^2 *and* ξ^2*) such that* $Y^1, Y^2 \in E$. *We assume that*

- $\xi^1 \leq \xi^2$ *and* $f^1 \leq f^2$ *a.s.*
- f^1 *convex with respect to* z *for all* $t \in [0,T]$ *and* $y \in \mathbb{R}$, *uniformly Lipschitz with respect to* y *and has the growth*

$$|f(t,y,z)| \leq \alpha(t) + \beta(y) + \frac{\gamma}{2}|z|^2, \quad \forall (t,y,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}^d;$$

with $\beta \geq 0$, $\gamma > 0$ *and* $(\alpha_t)_{t\in[0,T]}$ *a non-negative progressively measurable stochastic process which admits exponential moment of all orders.*

Then,

$$\forall t \in [0,T], \quad Y_t^1 \leq Y_t^2 \quad a.s.$$

If moreover $Y_0^1 = Y_0^2$, *then*

$$P\left(\xi^1 - \xi^2 = 0; \int_0^T (f^1 - f^2)(t, Y_t^1, Z_t^1)\, dt = 0\right) > 0.$$

Proof. See Briand and Hu (2008), Theorem 5. □

Remark 3.3.10. *In the case of bounded terminal condition, the stochastic integral of the process* Z *is a BMO-martingale, Proposition 2.3.8. This property can be used to derive the uniqueness of solution. When the terminal condition in not bounded, the process* $\int_0^\cdot Z_t\, dB_t$ *is not, in general, a BMO-martingale any more.*

Example 3.3.11. *Consider the BSDEs*

$$Y_t = B_1^2 - \int_t^1 Z_s\, dB_s, \quad t \in [0,1] \tag{3.3.10}$$

with generator $f \equiv 0$ *and unbounded terminal condition* B_1^2. *The random variable* B_1^2 *has exponential moments of all orders* $\lambda \in (0, \frac{1}{2})$. *By Theorem 3.3.4, existence of solution of Equation* (3.3.10) *is not guaranteed. We have* $Y_t = E\left[B_1^2 | \mathcal{F}_t\right]$ *for all* t. *Hence,*

$$\int_0^1 Z_s\, dB_s = B_1^2 - E\left[B_1^2\right].$$

Since $\exp\left(\frac{1}{2}B_1^2\right)$ *is not integrable,* $\mathcal{E}\left(\int_0^\cdot Z_s\, dB_s\right)$ *is not a uniformly integrable martingale. In follows from Theorem A.1.2 that the stochastic integral of* Z *is not a BMO-martingale.*

As a by-product of the latter theorem, we also obtain a uniqueness result.

Corollary 3.3.12. *Under the assumptions of Theorem 3.3.9, assume in addition that* ξ *admits exponential moments of all orders. Then, BSDE* (3.1.1) *has a unique solution* $(Y,Z) \in E \times H^p(\mathbb{R}^d)$ *for each* $p \geq 1$.

Proof. See Corollary 3.3.8 and Briand and Hu (2008), Corollary 4 for $(Y, Z) \in E \times H^p(\mathbb{R}^d)$ for each $p \geq 1$. □

Remark 3.3.13. *A noteworthy fact is that all the results on existence and uniqueness for quadratic BSDEs we have stated are applied to one-dimensional BSDEs, (i.e. Y is a one-dimensional process, $n = 1$). Even for a bounded terminal condition, no general results of existence or uniqueness for multi-dimensional quadratic BSDEs is available. Frei and Reis (2011) present the counterexample*[3] *of a 2-dimensional quadratic BSDE with bounded terminal condition but which does not have a solution. They justify this result by the fact that the dimension matters a lot in integration of stochastic processes.*

3.3.2.3 Stability

We restrict the study of stability of quadratic BSDEs to the case of bounded terminal wealth. Let us start by giving this important result, which is actually a generalisation of Proposition 2.4 of Kobylanski (2000), and is given as stated by Briand and Hu (2006).

Proposition 3.3.14 (Monotone stability). *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of $\mathcal{F}_T$-measurable bounded random variables and $(f_n)_{n\in\mathbb{N}}$ be a sequence of generators which are continuous with respect to (y, z). We assume that $(\xi_n)_{n\in\mathbb{N}}$ converges to ξ a.s., $(f_n)_{n\in\mathbb{N}}$ converges locally uniformly on $[0, T] \times \mathbb{R} \times \mathbb{R}^d$ to f, and in addition that the sequences $(\xi_n)_{n\in\mathbb{N}}$ and $(f_n)_{n\in\mathbb{N}}$ are such that $\sup_{n\in\mathbb{N}} \|\xi_n\|_\infty < +\infty$, and*

$$\sup_{n\in\mathbb{N}} |f_n(t, y, z)| \leq \alpha + \beta|y| + \frac{\gamma}{2}|z|^2,$$

with $\alpha, \beta \geq 0$, and $\gamma > 0$. If for each $n \geq 1$, the BSDE with parameters f_n and ξ_n has a solution (Y^n, Z^n) in $\mathcal{S}^\infty(\mathbb{R}) \times H^2(\mathbb{R}^d)$ such that $(Y^n)_{n\in\mathbb{N}}$ is non-decreasing (respectively, non-increasing), then a.s., $(Y^n)_{n\in\mathbb{N}}$ converges uniformly on $[0, T]$ to $Y = \sup_{n\in\mathbb{N}} Y^n$ (respectively, $Y = \inf_{n\in\mathbb{N}} Y^n$), $(Z^n)_{n\in\mathbb{N}}$ converges to some Z in $H^2(\mathbb{R}^d)$ and (Y, Z) is a solution in $L^\infty(\mathbb{R}) \times H^2(\mathbb{R}^d)$ to the BSDE with parameters f and ξ.

Proof. See Briand and Hu (2006), Lemma 3. □

Theorem 3.3.15 (Stability). *Let $(\xi_n)_{n\in\mathbb{N}}$ and $(f_n)_{n\in\mathbb{N}}$ be two sequences of parameters of BSDEs. We assume that there exist $\alpha, \beta, b \in \mathbb{R}$ and a non-decreasing function c such that for all $n \in \mathbb{N}$ the function f_n satisfies*

$$|f_n(t, y, z)| \leq \alpha + \beta|y| + c(|y|)|z|^2,$$

and for all n there exists (Y^n, Z^n) in $L^\infty(\mathbb{R}) \times H^2(\mathbb{R}^d)$ solution of the BSDE with parameters f_n and ξ_n. If $(f_n)_{n\in\mathbb{N}}$ converges to f locally uniformly on

[3] See Section (3.3.1).

$[0,T] \times \mathbb{R} \times \mathbb{R}^d$, *and if* $(\xi_n)_{n\in\mathbb{N}}$ *converges to* ξ *in* $L^\infty(\mathbb{R})$ *such that* f *and* ξ *satisfy the conditions of Theorem 3.3.7, then there exists a process* $(Y,Z) \in L^\infty(\mathbb{R}) \times H^2(\mathbb{R}^d)$ *such that the sequences* $(Y^n)_{n\in\mathbb{N}}$ *converges to* Y *uniformly on* $[0,T]$, $(Z^n)_{n\in\mathbb{N}}$ *converges to* Z *in* $H^2(\mathbb{R}^d)$ *and* (Y,Z) *is the solution of the BSDE with parameters* f *and* ξ.

Proof. See Kobylanski (2000), Theorem 2.8. □

Before assessing the topic of differentiability of quadratic BSDEs, let us open a short parenthesis to discuss the existence of solutions of BSDEs with jumps.

3.3.3 Quadratic BSDEs with Jumps

Backward stochastic differential equations with jumps arise in utility maximization problems in a discontinuous setting, i.e. when the price dynamic of the stock or the bond is a local martingale driven not only by a Brownian motion, but additionally by a Poisson point process. BSDEs with jumps take the form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, U_s)\, ds - \int_t^T Z_s\, dB_s - \int_t^T \int_{\mathbb{R}^*} U_s(q) \tilde{N}_k(ds, dq), \quad (3.3.11)$$

where the solution is a triple of predictable processes (Y,Z,U), and $\tilde{N}_k$ denotes the compensated Poisson random measure[4].

The issues of existence and uniqueness of solutions of this class of BSDEs (we consider here the cases of quadratic generator in z and bounded terminal condition) have recently been extensively studied. Among other works, we can quote Rong (2007), where an existence and uniqueness result is proved. In that work, the author first proves an existence result for a BSDE with Lipschitz generator and, by a change of variables of the form $Y_t = \frac{1}{X_t}$ where X is the first component of the solution of the Lipschitz BSDE, the existence result is obtained for the quadratic case. Of course, the technique requires Y to be non zero for every t. Ankirchner *et al.* (2010a) study the specific case of a BSDE with only one possible jump which occurs at a random time τ and with terminal condition of the form $\xi 1_{\tau>T} + \zeta 1_{\tau\le T}$. The solution of the BSDE with jumps is constructed thanks to the solution of two continuous BSDEs. Considering a particular generator, Morlais (2010) proves an existence result by extending the proof of Kobylanski (2000) to the case of a discontinuous BSDE. She works with the specific function

$$f(t,u,z) = \frac{1}{2}\eta\sigma_t^2 dist_t^2 \left(\frac{1}{\sigma}\left(\frac{\theta}{\eta} - z\right), C\right) - \frac{1}{2\eta}\theta_t^2 + \theta_t z + \inf_{\pi\in C} |\pi\gamma_t - u|_\eta,$$

[4]See Subsection 5.2.1.

where $|\pi\gamma_t - u|_\eta = \int_{\mathbb{R}^*} \left(\frac{1}{\eta}(\exp\big(\eta(u(x) - \pi\gamma_t)\big) - 1) - \pi\gamma_t + u(x)\right) n(dx)$, and n denotes the Lévy measure (which is not assumed to be finite). The function f satisfies the assumptions

(**HJ1**) For all $(z, u) \in \mathbb{R} \times (L^2 \cap L^\infty)(n)$,

$$-\theta_s z - \frac{|\theta_s|}{2\eta} \leq f(s, z, u) \leq \frac{\eta}{2}|z|^2 + |u|_\eta, \quad P \ a.s.$$

(**HJ2**) There exist C positive $\kappa \in BMO(P)$ and a process $\gamma = (\gamma_t(u, u'))_{t\in[0,T]}$ such that for all $(z, z') \in \mathbb{R} \times \mathbb{R}$, $u \in L^2(n)$,

$$|f(s, z, u) - f(s, z', u)| \leq C(\kappa_s + |z| + |z'|)|z - z'|$$

and

$$f(s, z, u) - f(s, z, u') \leq \int_{\mathbb{R}^*} \gamma_s(u, u')(x)(u(x) - u'(x))n(dx).$$

Theorem 3.3.16. *For any BSDE of the form* (3.3.11) *with generator f satisfying hypothesis (**HJ1**) and (**HJ2**) and with terminal condition ξ which is an arbitrary bounded $\mathcal{F}_T$-random variable, there exists a unique solution* $(Y, Z, U) \in \mathcal{S}^{\mathbb{R}} \times L^2(\mathbb{R}, P) \times L^2(\mathbb{R}^*, n)$ *to BSDE* (3.3.11).

Proof. See Morlais (2010), Section 3.3 and Lemma 4. □

3.4 Differentiability

In financial applications, the processes Y and Z in the BSDE along with the terminal condition ξ and the generator f generally depend on some parameters. The aim of this section is to discuss the properties of differentiability of the solution of quadratic BSDEs and FBSDEs with respect to a parameter and the differentiability in the Malliavin's sense. Most of the results of this section are taken from the Ph.D. thesis of Dos Reis, see Reis (2010). Since we also study the Malliavin's differentiability, we start by defining the Malliavin's derivative of a random variable.

Let $n \in \mathbb{N}$, $f \in C_b^\infty(\mathbb{R}^n)$ and F a random variable of the form $F = f(\int_0^T h_s^1 dB_s, \ldots, \int_0^T h_s^n dB_s)$, and $h^i \in L^2(0, T; \mathbb{R})$, $0 \leq i \leq n$. We define the derivative of F as the random variable

$$DF = \sum_{i=1}^n \partial_i f(\int_0^T h_s^1 dB_s, \ldots, \int_0^T h_s^n dB_s)h^i.$$

We can also define the second order derivative of F as $D^{(2)}F$, with

$$D^{(2)}F = \sum_{i,j=1}^n \partial_{ij}^2 f(\int_0^T h_s^1 dB_s, \ldots, \int_0^T h_s^n dB_s)h^i h^j.$$

In general, we define the k-th iterated derivative of F, $D^{(k)}F$ $k \in \mathbb{N}^*$. For $k, p \in \mathbb{N}^*$, we consider $\mathbb{D}^{k,p}(\Omega)$ the set of Malliavin's differentiable random variables defined as in Subsection 3.1.0.1. Note that $D^{(k)}$ is a closable operator, and we will use the notation $D^{(k)}$ to actually mean its closed extension (for more details see Nualar (1995)). Now, we give an important result that will be used in Subsection 3.4.2.

Lemma 3.4.1. *Let $(F_n)_{n\in\mathbb{N}^*}$ be a sequence of random variables in $\mathbb{D}^{1,2}(\Omega)$ which converges to F in $L^2(\Omega)$ and such that*

$$\sup_{n\in\mathbb{N}^*} \|DF_n\|_{L^2(\Omega\times[0,T])} < \infty.$$

Then F belongs to $\mathbb{D}^{1,2}(\Omega)$ and the sequence of derivatives $\{DF_n, n \in \mathbb{N}^*\}$ converges to DF in the weak topology of $L^2(\Omega \times [0,T])$.

Proof. See Nualar (1995) □

This lemma will give the key argument to go from the Malliavin's differentiability of Lipschitz BSDEs to Malliavin's differentiability of quadratic BSDEs. Let us consider the following quadratic BSDE with dependence on an Euclidean parameter $x \in \mathbb{R}$,

$$Y_t^x = \xi_x + \int_t^T f(s, x, Y_s^x, Z_s^x)\, ds - \int_t^T Z_s^x\, dB_s. \tag{3.4.1}$$

The concern is to study both the (classical) differentiability of the map $x \mapsto (Y^x, Z^x)$ with respect to the topology of the Banach spaces $\mathbb{R}$ and $\mathcal{S}^2(\mathbb{R}) \times H^2(\mathbb{R}^d)$, and the variational differentiability of the (non-parametrized) BSDE (3.1.1).

3.4.1 Classical Differentiability

To start with, let us give the main assumptions under which the differentiability result will be stated.

Let $\mathcal{O}$ be a non-empty subset of $\mathbb{R}$.

(**CD1**) The set $\mathcal{O}$ is open and convex, the function $f : \Omega\times[0,T]\times\mathcal{O}\times\mathbb{R}\times\mathbb{R}^d \mapsto \mathbb{R}$ is adapted and continuously differentiable with respect to (y, z) and the mapping $x \mapsto \xi_x$ from $\mathcal{O}$ to $L^\infty(\mathcal{O})$ is continuously differentiable, with $\Delta_x\xi_x \in L^2(\mathcal{O})$.

(**CD2**) There exist $M > 0$ and a non-negative process $(K_t(x))_{t\in[0,T]}$ such that f and ξ satisfy the following boundary conditions for all $(t, x, y, z) \in$

$[0,T] \times \mathcal{O} \times \mathbb{R} \times \mathbb{R}^d$,

$$
\begin{aligned}
|f(t,x,y,z)| &\leq M(1+|y|+|z|^2) \quad \text{a.s.},\\
|\Delta_x f(t,x,y,z)| &\leq K_t(x)(1+|y|+|z|^2) \quad \text{a.s.},\\
|\Delta_y f(t,x,y,z)| &\leq M \quad \text{a.s.},\\
|\Delta_z f(t,x,y,z)| &\leq M(1+|z|) \quad \text{a.s.},\\
\sup_{x\in\mathcal{O}} \|\xi_x\|_{L^\infty(\mathcal{O})} &< \infty,\\
\sup_{x\in\mathcal{O}} \|\Delta_x \xi_x\|_{L^2(\mathcal{O})} &< \infty.
\end{aligned}
$$

(**CD3**) Suppose that (CD1) and (CD2) hold, we further make the following assumptions. For $x \in \mathcal{O}$, $i \in \{1,\ldots,m\}$, $h \in \mathbb{R}^*$ and $e_i \in \mathbb{R}^m$, canonical unit vector, let $\nu^{x,h,i} = (\xi_{x+he_i} - \xi_x)/h$ with h such that $x + he_i \in \mathcal{O}$. For all $p \geq 1$, there exists $C > 0$ such that for all $x, x' \in \mathcal{O}$, $h, h' \in \mathbb{R}^*$ for which $(x + he_i) \in \mathcal{O}$, we have

$$E\left[|\nu^{x,h,i} - \nu^{x',h',i}|^2\right] \leq C(|x-x'|^2 + |h-h'|^2)^p.$$

The partial derivatives $\Delta_y f$ and $\Delta_x f$ satisfy the Lipschitz condition in (y,z) and the integrability condition

$$E\left[\left(\int_0^T |(\Delta_x f)(s,x,Y_s^x,Z_s^x) - (\Delta_x f)(s,x',Y_s^{x'},Z_s^{x'}|\, ds\right)^2\right] \leq C(|x-x'|^2)^p,$$

with $x, x' \in \mathcal{O}$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^d$ and (Y^x, Z^x) and $(Y^{x'}, Z^{x'})$, respectively solution of (3.4.1) with x and x'.

The differentiability result is stated as follows.

Theorem 3.4.2. *Let $x \in \mathcal{O}$, and (Y^x, Z^x) a solution of BSDE* (3.4.1). *If (CD1) and (CD2) hold, the function $\mathcal{O} \to \mathcal{S}^2(R) \times H^2(\mathbb{R}^d)$, $x \mapsto (Y^x, Z^x)$ is differentiable and its derivative $(\Delta Y^x, \Delta Z^x)$ solves the BSDE*

$$\Delta Y_t^x = \Delta\xi_x + \int_t^T \langle(\Delta f)(s,x,Y_s^x,Z_s^x),(1,\Delta Y_s^x,\Delta Z_s^x)\rangle\, ds - \int_t^T \Delta Z_s^x\, dB_s, \tag{3.4.2}$$

where $\Delta f = (\Delta_x f, \Delta_y f, \Delta_z f)'$.

If we further assume that (CD3) holds, then there exists a function $\Omega \times [0,T] \times \mathcal{O} \to \mathbb{R} \times \mathbb{R}^d$, $(\omega,t,x) \mapsto (Y_t^x(\omega), Z_t^x(\omega))$ such that for almost all ω, $t \mapsto Y_t^x$ is continuous and $x \mapsto Y_t^x$ is continuously differentiable with respect to x.

Proof. See Reis (2010), Theorem 3.1.3. $\square$

Remark 3.4.3. *1. Assumptions (CD1) and (CD2) are sufficient for existence and uniqueness of a solution of* (3.4.1) *(see Theorem 3.3.2 and Theorem 3.3.7). Therefore, the function $x \mapsto (Y^x, Z^x)$ is well defined.*

2. The second part of the theorem says, in other words, that the component Y^x of the solution of the BSDE is pathwise continuously differentiable. In particular, almost all its trajectories are continuous.

3. As Reis points out in his thesis, existence and uniqueness of solutions of Equation (3.4.1) *is not a priori guaranteed. The existence and uniqueness follow as a by-product of the theorem and its proof.*

One of the most important applications of the differentiability of such solutions of BSDEs in finance is the estimation of Greeks, or derivatives of the price of the contract with respect to the stock price or other market parameters.

3.4.2 Malliavin's Differentiability

In this paragraph, we assess the differentiability of solutions of quadratic BSDEs in a weak sense. We do not assume any topological structure on the space Ω. The result we are going to state roughly says that under some boundedness assumptions on the generator of the BSDE, the solution inherits the Malliavin's differentiability of the terminal condition. We start by giving the assumptions under which the differentiability holds.

Let $(D_u f(t,y,z))_{u,t\in[0,T]}$ be the Malliavin's derivative of f.

(MD1) The function $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is an adapted and continuously differentiable function with respect to (y,z) and there exists a constant $M \geq 0$ such that for all $(u,t,y,z) \in [0,T]^2 \times \mathbb{R} \times \mathbb{R}^d$,

$$\begin{aligned}
|f(t,y,z)| &\leq M(1+|y|+|z|^2) \quad \text{a.s.}\\
|\Delta_y f(t,y,z)| &\leq M \quad \text{a.s.}\\
|\Delta_z f(t,y,z)| &\leq M(1+z) \quad \text{a.s.}\\
|D_u f(t,y,z)| &\leq K_u^1(t)(1+|y|+|z|) + K_u^2(t)|z|^2 \quad \text{a.s.,}
\end{aligned}$$

where $(K_u^1(t))_{u,t\in[0,T]}$ and $(K_u^2(t))_{u,t\in[0,T]}$ are two processes satisfying

$$\int_0^T E\left[\|K_u^1(t)\|^2_{H^2(\mathbb{R})} + \|K_u^2(t)\|^2_{\mathcal{S}^2(\mathbb{R})}\right] du < \infty.$$

(**MD2**) The random variable $\xi \in L^\infty(\mathbb{R})$ is Malliavin differentiable and belongs to $\mathbb{D}^{k,p}(\mathbb{R})$.

We now give the (first order) Malliavin's differentiability result for quadratic BSDEs.

Theorem 3.4.4. *Let (Y, Z) a solution of BSDE (3.1.1). If (MD1) and (MD2) hold, then (a version of) $(D_u Y_t, D_u Z_t)_{u,t\in[0,T]}$ is the unique solution of the following BSDE: for $t \in [u, T]$,*

$$\begin{aligned} D_u Y_t &= D_u\xi - \int_t^T D_u Z_s \, dB_s \\ &\quad + \int_t^T \left[D_u f(s, Y_s, Z_s) + \langle (\Delta f)(s, Y_s, Z_s), D_u(Y_s, Z_s) \rangle \right] ds, \qquad (3.4.3) \\ D_u Y_t &= 0 \quad \text{and} \quad D_u Z_t = 0, \quad t \in [0, u). \end{aligned}$$

Sketch of the proof. The reader may find the detailed proof in Reis (2010).

The main arguments of the proof are Lemma 3.4.1 and the Malliavin's differentiability of Lipschitz BSDEs. The technique consists of constructing a sequence of Lipschitz functions $(f_n)_{n\geq 1}$ that converges to the quadratic driver f, applying the differentiability result to each Lipschitz BSDE with parameters f_n and ξ, then passing to the limit by means of Lemma 3.4.1.

For $n \in \mathbb{N}$, define $\tilde{h}_n : \mathbb{R} \to \mathbb{R}$ to be a continuously differentiable function which converges locally uniformly to the identity such that $\tilde{h}_n(z) = n + 1$ for $z > n+2$, $\tilde{h}_n(z) = z$ for $|z| \leq n$, $\tilde{h}_n(z) = -(n+1)$ for $z < -(n+2)$ and satisfies the boundedness conditions $|\tilde{h}_n(z)| \leq |z|$; $|\tilde{h}_n(z)| \leq n + 1$ and $|\Delta\tilde{h}_n(z)| \leq 1$ for all $n \geq 1$ and $z \in \mathbb{R}$. In addition, $(\Delta\tilde{h}_n)_n$ converges locally uniformly to 1. Further define the function $h_n : \mathbb{R}^d \to \mathbb{R}^d$, $z \mapsto (\tilde{h}_n(z_1), \ldots, \tilde{h}_n(z_d))$ and $f_n : \Omega \times [0, T] \times \mathbb{R}^d \to \mathbb{R}$, $(t, y, z) \mapsto f_n(t, y, z) = f(t, y, h_n(z))$. Since f is continuous and $(\tilde{h}_n)_{n\in\mathbb{N}}$ converges locally uniformly to the identity, the sequence $(f_n)_{n\in\mathbb{N}}$ converges locally uniformly to f. We want to show that for each n, f_n is Lipschitz continuous in (y, z). Let $n \in \mathbb{N}$ be fixed. Let $(t, y, z), (t, y', z') \in [0, t] \times \mathbb{R} \times \mathbb{R}^d$, by the mean value theorem we have,

$$\begin{aligned} |f_n(t, y, z) - f_n(t, y', z')| \leq &\max_{[0,T]\times\mathbb{R}\times\mathbb{R}^d} |\Delta_y f_n(\cdot, \cdot, \cdot)||y - y'| \\ &+ \max_{[0,T]\times\mathbb{R}\times\mathbb{R}^d} |\Delta_z f_n(\cdot, \cdot, \cdot)||z - z'|, \quad a.s. \end{aligned}$$

On the first hand, we have from assumption (MD1) $\max_{[0,T]\times\mathbb{R}\times\mathbb{R}^d} |\Delta_y f_n(\cdot, \cdot, \cdot)| \leq M$. On the other hand, for all $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$ we have

$$\begin{aligned} |\Delta_z f_n(t, y, z)| &= |\Delta_z f(t, y, h_n(z))\Delta_z h_n(z)| \qquad (3.4.4) \\ &\leq M(1 + |h_n(z)|) \qquad (3.4.5) \\ &\leq M(2 + n), \qquad (3.4.6) \end{aligned}$$

where (3.4.4) comes from the chain rule, (3.4.5) is a consequence of (MD1) and the fact that $|\Delta_z h_n(z)| \leq 1$, and (3.4.6) follows from the condition $|\tilde{h}_n| \leq n+1$. This shows that f_n satisfies the Lipschitz condition in the spatial variables for all n, and is continuous as the composition of two continuous functions. Thus, for $n \in \mathbb{N}$ consider the BSDE

$$Y_t^n = \xi + \int_t^T f_n(s, Y_s^n, Z_s^n)\, ds - \int_t^T Z_s^n \, dB_s, \quad t \in [0, T]. \qquad (3.4.7)$$

By Reis (2010), Lemma 3.2.2, (Y^n, Z^n) is Malliavin differentiable for all n, and a version of $(D_u Y_t^n, D_u Z_t^n)_{u,t\in[0,T]}$ is solution a of the following BSDE: for $t \in [u, T]$

$$\begin{aligned} D_u Y_t^n = D_u \xi - \int_t^T D_u Z_s^n \, dB_s \\ + \int_t^T [D_u f_n(s, Y_s^n, Z_s^n) + \langle (\Delta f_n)(s, Y_s^n, Z_s^n), D_u(Y_s^n, Z_s^n)\rangle] \, ds; \end{aligned} \tag{3.4.8}$$

$$D_u Y_t^n = 0 \quad \text{and} \quad D_u Z_t^n = 0, \quad t \in [0, u).$$

Moreover, for all $u \in [0, T]$, $\|D_u Y^n\|_{L^2(\Omega\times[0,T])}$ and $\|D_u Z^n\|_{L^2(\Omega\times[0,T])}$ are uniformly bounded with respect to n. The rest of the proof consists in establishing the (weak) convergence term by term of Equation (3.4.8) to Equation (3.4.3). By Reis (2010) Lemma 3.2.1, $(Y^n)_{n\in\mathbb{N}}$ converges to Y in $L^2(\Omega)$. Lemma 3.4.1 yields that there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $(D_u Y^{n_k})_{k\in\mathbb{N}}$ converges to $D_u Y$ weakly in $L^2(\Omega \times [0, T])$. The convergence of the second term of the right hand side of (3.4.8) is a consequence of the Itô isometry property of the stochastic integral, and the convergence of the last term comes from the dominated convergence theorem. This concludes the proof of the theorem. $\square$

As a by-product of the latter theorem and its proof, we have the following result.

Corollary 3.4.5. *Under the assumptions of Theorem 3.4.4, Equation* (3.4.3) *admits at least one solution.*

One of the most important applications of the Malliavin's differentiability of quadratic BSDEs comes from the fact that the control process can be represented has the Malliavin's derivative of the process Y. This enables to derive an explicit formula of the optimal hedge in terms of the derivative of the indifference price, see Section 5.1 and Ankirchner *et al.* (2010*b*).

3.5 Applications of Quadratic BSDEs

Quadratic BSDEs have found a strong foothold in several subjects in Mathematics. We conclude this chapter by giving two areas of application different from stochastic control; namely, non-linear PDEs and behavioural finance.

3.5.1 Non-Linear Feynman-Kac Formula

BSDEs are useful to derive a representation of the solution of some second order linear PDEs, as discussed in Subsection 3.1.1 in the case of linear BSDEs. In this subsection, we discuss an application of BSDEs to PDEs which are quadratic with respect to the gradient of the solution. Let $(t, x, y, z) \mapsto f(t, x, y, z)$ be a function with quadratic growth in z and Lipschitz in y. Consider the diffusion defined by Equation (3.1.4). We want to solve the PDE

$$\partial_t v(t,x) + \mathcal{L}v(t,x) + f(t,x,v(t,x),\sigma'\nabla_x v(t,x)) = 0, \quad v(T,.) = g. \tag{3.5.1}$$

With the operator $\mathcal{L}$, generator of (3.1.4), defined as

$$\begin{aligned}\mathcal{L}u &= \sum_{i=1}^{n} b_i(t,x)\frac{\partial u}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} a_{i,j}(t,x)\frac{\partial^2 u}{\partial x_i \partial x_j} \quad (\text{with} \quad a = \sigma\sigma') \\ &= (D_x u)' b(t,x) + \frac{1}{2} tr\big((D_{xx}u)a(t,x)\big),\end{aligned}$$

for any function u smooth enough, and g a bounded function. We want to give a probabilistic representation for the solution of the PDE when f is quadratic. This is the result of the following proposition given as stated in El Karoui *et al.* (1997), Proposition 4.3.

Proposition 3.5.1. *Let v be a function of class $C^{1,2}$ and suppose that there exists a constant C such that*

$$|v(s,x)| + |\sigma(s,x)'\nabla_x v(s,x)| \le C(1+|x|), \quad \forall (s,x) \in [0,T]\times\mathbb{R}^n.$$

Moreover, assume v to be solution of the quasilinear parabolic PDE (3.5.1).

Then, we have $Y_t^{u,x} = v(t, X_t^{u,x})$ and $Z_t^{u,x} = \sigma(t, X_t^{u,x})'\nabla v(t, X_t^{u,x})$, $t \in [u,T]$, where $(X^{u,x}, Y^{u,x}, Z^{u,x})$ is the solution of (3.1.4)-(3.1.5).

Proof. Since v is smooth enough, Itô's formula yields

$$\begin{aligned} dv(t,X_t^{u,x}) &= \partial_t v(t,X_t^{u,x})\,dt + (\nabla_x v(t,X_t^{u,x}))' b(t,X_t^{u,x})\,dt \\ &+ (\nabla_x v(t,X_t^{u,x}))'\sigma(t,X_t^{u,x})\,dB_t + \frac{1}{2} tr(D_{xx}v(t,X_t^{u,x})\sigma(t,X_t^{u,x})'\sigma(t,X_t^{u,x}))\,dt.\end{aligned}$$

This implies, for $t \in [u,T]$

$$\begin{aligned} v(t,X_t^{u,x}) &= v(T,X_T^{u,x}) - \int_t^T \partial_t v(s,X_s^{u,x})\,ds - \int_t^T (\nabla_x v(s,X_s^{u,x}))' b(s,X_s^{u,x})\,ds \\ &- \int_t^T \frac{1}{2} tr(D_{xx}v(s,X_s^{u,x})a(s,X_s^{u,x}))\,ds - \int_t^T (\nabla_x v(s,X_s^{u,x}))'\sigma(s,X_s^{u,x})\,dB_s \\ &= v(T,X_T^{u,x}) - \int_t^T \mathcal{L}v(s,X_s^{u,x})\,ds - \int_t^T (\nabla_x v(s,X_s^{u,x}))'\sigma(s,X_s^{u,x})\,dB_s.\end{aligned}$$

Since v solves Equation (3.5.1), we have the BSDE

$$\begin{aligned} v(t,X_t^{u,x}) = g(X_T^{u,x}) &+ \int_t^T f(s,x,v(s,X_s^{u,x}),\sigma'\nabla v(s,X_s^{u,x}))\,ds \\ &- \int_t^T (\nabla v(s,X_s^{u,x}))'\sigma(s,X_s^{u,x})\,dB_s,\end{aligned}$$

which completes the proof thanks to the results on existence and uniqueness of quadratic BSDEs. In particular, v is defined for all t, x by $v(t,x) = Y_t^{t,x}$. $\square$

It is worth noticing that if v is not assumed to be smooth enough, v is not a classical solution of (3.5.1), but the function $v(t, x) = Y_t^{t,x}$ is a viscosity solution of (3.5.1), see Briand and Hu (2008).

It is rather interesting to point out that the Feynman-Kac's formula gives a kind of dual relationship between solutions of BSDEs and those of PDEs. In fact, if the solution of the PDE is assumed to be known, then the representation of the latter proposition gives a solution of BSDE (3.1.5), and therefore a solution of FBSDE (3.1.5)-(3.1.4). On the other hand, if we instead assume to know the solution of (3.1.5)-(3.1.4), then the representation gives the solution, or at least the viscosity solution, of the parabolic PDE (3.5.1). Another application of the formula is the numerical approximation of BSDEs. A link has also been made between BSDEs and classical solutions of a class of PDEs, see for instance Ma *et al.* (1994) where the authors establish a four-step scheme for approximation of BSDEs via solutions of PDEs.

3.5.2 Generalised Stochastic Differential Utility

Quadratic BSDEs find an important application in the theory of decision-making in Economy and Finance. They enable to define a class of dynamic utility functions which use the information available. In fact, if we consider a model with consumption rate process $(c_t)_{t\in[0,T]}$ the utility U_t at time t can be termed as solution of the recursive equation

$$\frac{dU_t(c_t)}{dt} = -f(t, c_t, U_t(c_t), \omega),$$

where the stochastic function f is the intertemporal aggregator. We put U the utility from the terminal consumption. Since c is an adapted process, we replace U by its best adapted approximation

$$U_t(c_t) = E\left[U(c_T) + \int_t^T f(s, c_s, U_s(c_s))\, ds \;\Big|\; \mathcal{F}_t\right], \quad t \in [0, T].$$

This last formula defines the stochastic differential utility (SDU) as introduced by Duffie and Epstein, and led the authors to an independent (from Pardoux and Peng) formulation of BSDEs. Note the conditioning on $\mathcal{F}_t$ in the definition of U_t, which implies that the investor knows what has happened up to time t, and uses that information to define the (dynamic) utility. Using the fact that the process defined above is a martingale, Lazrak and Quenez (2003) derived a generalized SDU (GSDU) thanks to the martingale representation theorem. Indeed, there exists a process $(Z_t)_{t\in[0,T]}$ such that

$$U_t(c_t) = U(c_T) + \int_t^T f(s, c_s, U_s(c_s))\, ds - \int_t^T Z_s\, dB_s, \quad t \in [0, T].$$

For the sake of generality, we allow f to depend on Z and define the GSDU U for all t by

$$U_t(c_t) = U(c_T) + \int_t^T f(s, c_s, U_s(c_s), Z_s)\, ds - \int_t^T Z_s\, dB_s. \tag{3.5.2}$$

In the case where f takes the form $f(t, x, u, z) = \log(c) - \beta y - (\alpha/2)z^2$, the GSDU defined by (3.5.2) yields a solution criterion to some problems related to robust control theory, see Lazrak and Quenez (2003).

Chapter 4

Measure Solutions of BSDEs

The seminal works of Black, Scholes and Merton in the 1970s have provided a classical solution to the problem of pricing and hedging of an European contingent claim in a complete financial market. The price is obtained by taking a conditional expectation with respect to a martingale measure and, the hedging strategy comes from the martingale representation theorem. The concept of measure solutions was introduced by Ankirchner, Imkeller and Popier in Ankirchner *et al.* (2009), to replicate the Black-Scholes-Merton machinery in the context of non-linear BSDEs.

4.1 Definition and Concept

We consider the BSDE

$$Y_t = \xi + \int_t^T f(s, Z_s)\, ds - \int_t^T Z_s\, dB_s, \quad t \in [0, T] \tag{4.1.1}$$

written on a probability space $(\Omega, \mathcal{F}, P)$ which carries a d-dimensional Brownian motion $(B_t)_{t\in[0,T]}$, and is endowed with the augmented Brownian filtration $\mathbb{F}$. We also assume f to be adapted with quadratic growth in the control variable.

By the fundamental theorem of asset pricing in a complete market, the non-arbitrage hypothesis leads to the existence of an equivalent probability measure under which the discounted price process is a martingale. This martingale measure eliminates the drift part of the process through a Girsanov change of probability whence, from the real world measure, we switch to a new measure. Analogously, can we find a probability measure that eliminates the generator of the BSDE? In other words, does there exist a probability Q that transforms BSDE (4.1.1) into a BSDE of the form

$$Y_t = \xi - \int_t^T Z_s\, dB_s^Q, \quad t \in [0, T], \tag{4.1.2}$$

where B^Q is a Brownian motion under Q? This transformation can be carried out in the following way:

From Equation (4.1.1), we factorize with respect to Z_s and we are led to

$$Y_t = \xi - \int_t^T \langle Z_s, [dB_s - g(s, Z_s)\, ds] \rangle, \quad t \in [0, T];$$

with $g : \Omega \times [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, a function that is continuous in z and satisfies the identity[1] $\langle z, g(t, z) \rangle = f(t, z)$ for all $(t, z) \in [0, T] \times \mathbb{R}^d$. Put $M_t = \int_0^t g(s, Z_s)\, dB_s$, $t \in [0, T]$. Assuming $\int_0^t \|g(s, Z_s)\|^2\, ds < \infty$ and $\mathcal{E}(M)$ to be a uniformly integrable martingale, $Q = \mathcal{E}(M) \cdot P$ is a probability measure and, by Girsanov's theorem, B^Q, defined by $B_t^Q = B_t - \int_0^t g(s, Z_s)\, ds$, is a Brownian motion under Q. Hence Equation (4.1.2) follows. Note that such a transformation will indeed ease the theory and as explained in Section 3.1, the component Z of the solution of (4.1.2) is obtained as a direct consequence of the martingale representation theorem, and the component Y is given by $Y_t = E^Q [\xi | \mathcal{F}_t]$.

These observations lead to the following definitions:

Definition 4.1.1. *Given BSDE* (4.1.1)*, we shall call sub-generator a function $g : \Omega \times [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying $\langle z, g(t, z) \rangle = f(t, z)$, and continuous for all $t \in [0, T]$.*

Definition 4.1.2. *A triplet (Y, Z, Q) is called a measure solution of the BSDE*[2] (4.1.1) *if f admits a sub-generator g, Q is a probability measure on $(\Omega, \mathcal{F})$, and (Y, Z) is a pair of $\mathbb{F}$-predictable stochastic processes such that $\int_0^T Z_s^2\, ds < \infty$, Q-a.s. and the following conditions are satisfied:*

$$\begin{aligned}
&B^Q = B - \int_0^{\cdot} g(s, Z_s)\, ds \quad \textit{is a } Q\textit{-Brownian motion},\\
&\xi \in L^1(\Omega, \mathcal{F}, Q),\\
&Y_t = E^Q [\xi | \mathcal{F}_t] = \xi - \int_t^T Z_s\, dB_s^Q, \quad t \in [0, T].
\end{aligned}$$

We assume the generator f to satisfy the following assumptions:

- (MS1) f is adapted with quadratic growth in the control variable
- (MS2) f admits a sub-generator such that there exists $\epsilon > 0$ and a predictable process Φ such that $\int_0^{\cdot} \Phi_s\, dB_s$ is a BMO-martingale and for every $|z| \le \epsilon$, we have $|g(t, z)| \le \Phi_t$.

[1] Henceforth, if there is no ambiguities, the scalar product $\langle ., . \rangle$ is merely denoted by $\langle a, b \rangle = ab$.

[2] We consider a generator f that does not depend on Y because if the generator depends on Y, so does the probability Q. Hence the formula $Y_t = E^Q [\xi | \mathcal{F}_t]$ does not give the solution any more.

In our setting of a multi-dimensional Brownian motion, we restrict the domain of the generator f to the set $\Omega \times [0,T] \times \mathcal{B}(O,\epsilon)$, where $\mathcal{B}(O,\epsilon)$ is the open ball of center O and radius ϵ.

Remark 4.1.3. *The choice of defining the generator for controls in $\mathcal{B}(O,\epsilon)$ can seem a bit restrictive in the sense that the theory we are describing will be valid only for controls not exceeding a certain value. However, this should be good enough for financial applications because the total amount that the trader is assumed to spend is finite and small enough not to influence the market. It is hence realistic to assume Z to be uniformly bounded.*

Remark 4.1.4. *It is far from being straightforward that $Q = \mathcal{E}(M) \cdot P$ is a probability measure. Ankirchner* et al. *(2009) state in their introduction that the usual Novikov's criterion fails to be applied. Assumption (MS2) provides the key argument to prove that Q is a probability.*

In the following example, we illustrate how measure solutions arise in a pricing problem in the Black-Scholes economy. The example is based on the complete market model of Subsection 3.1.1.

Example 4.1.5. *For simplicity, we consider the model with a single risky asset. We recall that the investor willing to buy the European claim ξ at time T constructs a portfolio with values following the dynamics*

$$dV_t = (r_t V_t + \pi_t \sigma_t \theta_t - c_t)\,dt + \pi_t \sigma_t\,dB_t. \tag{4.1.3}$$

The price of the claim is the time-zero value of the process V solving the equation above, with terminal condition $V_T = \xi$. Thus we have

$$V_t = \xi - \int_t^T \pi_s \sigma_s\,dB_s^Q, \quad t \in [0,T]$$

with

$$dB_t^Q = \left(\theta_t + \frac{r_t V_t - c_t}{\pi_t \sigma_t}\right) dt + dB_t \quad \textit{and} \quad \frac{dQ}{dP} = \mathcal{E}\left(\int_0^{\cdot} (\theta_t + \frac{r_t V_t - c_t}{\pi_t \sigma_t}) dB_t\right)_T.$$

The process B^Q is a Q-Brownian motion since all processes are bounded. If we consider the case $c_t = r_t = 0$, then Q becomes the risk-neutral probability. The triplet $(V, \sigma\pi, Q)$ is measure solution of (4.1.3)*. The price of the claim is $V_0 = E^Q[\xi]$, which is the formula obtained by the Black-Scholes mechanism.*

Beyond the construction of this new form of solutions of BSDEs, an interesting question that needs to be addressed is the relationship between the solution (Y,Z) of (4.1.1) (or strong solution of (4.1.1)) and the solution (Y,Z,Q) of (4.1.2) (or weak solution of (4.1.1)). This could help to derive many results on BSDEs in a rather easier way.

4.2 Link Between Measure Solutions and Strong Solutions

In this section, we discuss the link between the existence and the uniqueness of strong solutions and that of measure solutions. In the spirit of Ankirchner *et al.* (2009) and Briand and Hu (2008), we will consider both the cases of bounded and unbounded terminal conditions.

The result below, taken from Ankirchner *et al.* (2009), provides an equivalence property between existence of strong solutions and existence of measure solutions.

Theorem 4.2.1. *Assume that* ξ *is bounded, and* f *satisfies assumptions (MS1) and (MS2). Then* (Y, Z) *is a classical solution of BSDE* (4.1.1) *if, and only if, there exists a probability measure* Q *equivalent to* P *such that* (Y, Z, Q) *is a measure solution of BSDE* (4.1.1)*.*

Proof. It is rather easy to see that if (Y, Z, Q) is a measure solution, then (Y, Z) is a strong solution. Of course, let (Y, Z, Q) be a measure solution of BSDE (4.1.1). Then, Q is a probability measure under which $B^Q = \int_0^\cdot g(s, Z_s)\, ds - B$ is a Brownian motion, and (Y, Z) is given by

$$\begin{aligned}
Y_t = E^Q\left[\xi|\mathcal{F}_t\right] &= \xi - \int_t^T Z_s\, dB_s^Q \\
&= \xi + \int_t^T \langle Z_s, g(s, Z_s)\rangle\, ds - \int_t^T Z_s\, dB_s \\
&= \xi + \int_t^T f(s, Z_s)\, ds - \int_t^T Z_s\, dB_s.
\end{aligned}$$

Thus, (Y, Z) is a strong solution.

For the proof of the converse, we refer to Ankirchner *et al.* (2009), Theorem 1.1. One may also refer to the proof of our Theorem 4.2.5 where we use a reasoning similar to that of Ankirchner *et al.* (2009) to extend the result to a class of BSDEs with jumps. □

Remark 4.2.2. *Theorem 4.2.1 can be understood as follows: In the case of a bounded terminal condition, if the generator is continuous and satisfies the quadratic growth condition, then the existence of a strong solution is equivalent to the existence of a measure solution. For this reason, the existence results given in Chapter 3 can be recovered by means of measure solutions.*

A result like that of Theorem 4.2.1 is not available in the case of unbounded terminal conditions. We are going, in the sequel of this section, to give a counterexample which exhibits a BSDE having a strong solution which is not a measure solution. The counterexample is actually a particular case of the more general study of Ankirchner *et al.* (2009), Subsection 2.2. We consider the one-dimensional BSDE

$$Y_{t\wedge\tau} = \xi - \int_t^\tau Z_s \, dB_s + \int_t^\tau \frac{1}{2} Z_s^2 \, ds. \tag{4.2.1}$$

With $\tau = \inf\{t \geq 0 : B_t \leq t - 1\}$, $\xi = -4(\tau + 1)$ and B is one-dimensional. Thus the generator is given by the $\mathbb{R}$-valued function $f : (t, y, z) \mapsto \frac{1}{2}z^2$. The pair (Y, Z) defined by $Y_t = 4B_{t\wedge\tau} - 8t \wedge \tau$ and $Z = 4 \times 1_{[0,\tau]}$ is the solution of Equation (4.2.1). In fact,

$$\begin{aligned} \xi - \int_t^\tau Z_s dB_s + \int_t^\tau \frac{1}{2} Z_s^2 \, ds &= -4(\tau + 1) + 4(B_{t\wedge\tau} - B_\tau) - 8(\tau \wedge t - \tau) \\ &= 4B_{t\wedge\tau} - 8(\tau \wedge t), \end{aligned}$$

by definition of τ. Now we are going to show that the solution (Y, Z) defined above is not a measure solution. It suffices to show that $\mathcal{E}(M)$, with $M = \int_0^\cdot \frac{f(s,Z_s)}{Z_s} \, dB_s$, is not a martingale[3]. Because $\mathcal{E}(M)$ is a positive supermartingale with $\mathcal{E}(M)_0 = 1$, we have $E[\mathcal{E}(M)_t] \leq 1$ for all t. Thus, $\mathcal{E}(M)$ is a martingale if, and only if, $E[\mathcal{E}(M)_t] = 1$ for every t.

$$\begin{aligned} E[\mathcal{E}(M)_\tau] &= E\left[\exp\left(\int_0^\tau \frac{1}{2} Z_s \, dB_s - \frac{1}{8}\int_0^\tau Z_s^2 \, ds\right)\right] \\ &= E[\exp(2(B_\tau - \tau))] \quad \text{by definition of} \quad Z \\ &= E[\exp(2(\tau - 1 - \tau))] \quad \text{by definition of } \tau \\ &< 1. \end{aligned}$$

Hence (Y, Z) is a strong solution but not a measure solution of (4.2.1).

It is noteworthy to point out that the terminal condition ξ of Equation (4.2.1) does not admit exponential moments of all orders. In fact, for $\lambda \in \mathbb{R}$, by the Laplace transform of τ, see for instance Revuz and Yor (1999), we have

$$E[\exp(\lambda\tau)] = \exp(-\sqrt{1 - 2\lambda} - 1).$$

Thus, the exponential moment exists only if $\lambda \leq \frac{1}{2}$. According to Corollary 3.3.12, Equation (4.2.1) could have several solutions.

Remark 4.2.3. *Equation* (4.2.1) *is a BSDE with random terminal time and quadratic generator. We do not study this type of BSDE in the present thesis. For more details, we refer to Makasu (2009) where the author characterises the exit time of solutions and provides a uniqueness result.*

---

[3]Therefore, Girsanov's theorem will not be applicable. One can also use Novikov's criterion to conclude.

4.2.1 Case of BSDEs with Jumps

In this subsection, we define the concept of measure solutions for a class of quadratic BSDEs with jumps and prove an analogous theorem to Theorem 4.2.1. We model the jump process by a compound Poisson process. More precisely, the jump times are modelled by a sequence $(T_n)_{n\geq 1}$ of (non-negative) random variables defined for $n \geq 1$ by $T_n = \sum_{i=1}^{n} \tau_i$ where $(\tau_i)_{i\geq 1}$ are independent exponential random variables. For a given $t \in [0, T]$, $N_t = \sum_{n\geq 1} \mathbf{1}_{t\geq T_n}$ is the number of jumps that have occurred up to time t. Since $(N_t)_{t\in[0,T]}$ is a Poisson process (with intensity ν_t), it is well-known that there are, almost surely, finitely many jumps. In our setting, we assume that the jump times are ordered, i.e. $T_1 \leq \cdots \leq T_n$. Since we will use the martingale representation theorem, the suitable filtration should be generated by both the Brownian motion and the Poisson process.

We consider a σ-finite measure α on the measurable space $(E, \mathcal{B}(E))$, with $E = [0, T] \times \mathbb{R}^*$. For simplicity, we will assume α to be the Lebesgue measure. Define the jump's intensity process by $(U_t(q))_{t\in[0,T]}$. Denote by p the Poisson random measure generated by the Poisson process, and its compensated analogue by

$$\tilde{p}(A) = p(A) - \alpha(A), \quad A \in \mathcal{B}(E),$$

that is assumed to take the form $\tilde{p}(ds, dq) = n(dq)\, ds$, with n the Lévy measure.

The jump process $(J_t^U)_{t\in[0,T]}$ is defined, for each $t \in [0, T]$, by

$$J_t = \int_0^t \int_{\mathbb{R}^*} U_s(q)\tilde{p}(ds, dq) = \int_0^t \int_{\mathbb{R}^*} U_s(q)(p(ds, dq) - \alpha(ds, dq)),$$

where for the sake of notational simplicity we omit the superscript U. The process $(J_t)_{t\in[0,T]}$ is a jump process whose jumps happen at random times T_n and have the intensity $U_{T_n}(q)$. Moreover, see Cont and Tankov (2004) Proposition 2.16 page 60, it is a martingale. Note that the random measure p is the derivative[4] (in the sense of distributions) of the Poisson process.

We consider the following form of BSDEs already defined in Subsection 3.3.3:

$$Y_t = \xi + \int_t^T f(s, Z_s, U_s)\, ds - \int_t^T Z_s\, dB_s - \int_t^T \int_{\mathbb{R}^*} U_s(q)\tilde{p}(ds, dq). \quad (4.2.2)$$

That is,

$$dY_t = -f(t, Z_t, U_t)\, dt + Z_t\, dB_t + U_t\, d\tilde{N}_t.$$

Hence, Y_t can be seen as a jump-diffusion process. By an extension of the Girsanov change of measure, see Runggaldier (2003), Theorem 2.5 and Section 4.1, there exist two processes θ and ϕ such that $dB_t^Q = dB_t - \theta_t\, dt$ and

[4] We will sometimes write $dJ_t = U_t d\tilde{N}_t$, where $\tilde{N}$ is the compensated Poisson process.

$d\tilde{N}_t^Q = d\tilde{N}_t - \phi_t \nu_t \, dt$ are, respectively, a Q-Brownian motion and a Q-Poisson martingale. The Random-Nikodym derivative of the measure Q is given by

$$V_t = \exp\left\{ \int_0^t \left[\phi_s \nu_s - \frac{1}{2}\theta_s^2 \right] ds + \int_0^t \theta_s \, dB_s + \int_0^t (1 - \nu_s \phi_s) d\tilde{N}_s \right\} \prod_{n=1}^{N_t} \phi_{T_n}.$$

Inserting B^Q and $\tilde{N}^Q$ in the BSDE leads to

$$dY_t = \big(-f(t, Z_t, U_t) + Z_t \theta_t + \nu_t(\phi_t - 1) U_t \big) \, dt + Z_t \, dB_t^Q + U_t \, d\tilde{N}_t^Q.$$

The measure Q is a martingale measure if θ and ϕ are chosen such that for all $t \in [0, T]$, $-f(t, Z_t, U_t) + Z_t \theta_t + \nu_t(\phi_t - 1)U_t = 0$. Hence, we take

$$\theta_t = \frac{f(t, Z_t, U_t)}{Z_t} - \frac{\nu_t(\phi_t - 1)U_t}{Z_t}, \quad \phi_t > 1 \quad \text{and} \quad t \in [0, T].$$

In addition, we take ϕ such that the process $\left(\frac{\nu_t(\phi_t - 1)U_t}{Z_t} \right)_{t \in [0,T]}$ is bounded. A simple example of such a process would be $\phi_t = |\frac{Z_t}{U_t}| + 1$, $t \in [0, T]$. With that choice of θ and ϕ, the BSDE becomes

$$dY_t = Z_t dB_t^Q + U_t d\tilde{N}_t^Q,$$

this implies $Y_t = E^Q[\xi | \mathcal{F}_t]$, and the processes Z and U are given by the extension of the martingale representation theorem to jump-diffusion processes, see Runggaldier (2003), Theorem 2.3. The reasoning is formalized in the following definition:

Definition 4.2.4. *Consider the filtration*

$$\mathcal{F}_t = \sigma \left\{ B_s, N_s, A : 0 \le s \le t, A \in \mathcal{N} \right\},$$

with $\mathcal{N}$ the collection of P-null sets from $\mathcal{F}$. A quadruplet (Y, Z, U, Q) is called a measure solution of the BSDE (4.2.2) if f admits a sub-generator, Q is a probability measure on $(\Omega, \mathcal{F})$, (Y, Z, U) is a triplet of $\mathbb{F}$-predictable stochastic processes such that Z is Q a.s. square integrable and U is Q a.s. $\alpha(dq, .)$-integrable and the following conditions hold:

$$\begin{aligned}
dB_t^Q &= dB_t - \frac{f(t, Z_t, U_t)}{Z_t} + \frac{\nu_t(\phi_t - 1)U_t}{Z_t}, \\
d\tilde{N}_t^Q &= d\tilde{N}_t - \phi_t \nu_t dt, \\
\xi &\in L^1(\Omega, \mathcal{F}, Q), \\
Y_t &= E^Q[\xi | \mathcal{F}_t] = \xi - \int_t^T Z_s \, dB_t^Q - \int_t^T U_s \, d\tilde{N}_t^Q.
\end{aligned}$$

Analogously to the case of BSDEs driven by Brownian motion, we have the following result:

Theorem 4.2.5. *Assume that ξ is bounded and f satisfies assumptions (MS1) and (MS2). Then (Y, Z, U) is a classical solution of BSDE* (4.2.2) *if, and only if, there exist probability measures Q equivalent to P such that (Y, Z, U, Q) are measure solutions of BSDE* (4.2.2)*.*

Proof. Let (Y, Z, U) be a strong solution of BSDE (4.2.2). Put

$$M = \int_0^{\cdot} \frac{f(s, Z_s, U_s)}{Z_s} - \frac{\nu_s(\phi_s - 1)U_s}{Z_s}\, dB_s,$$

such that

$$V = \exp\left(M - \frac{1}{2}[M]\right) \exp\left\{\int_0^{\cdot} \phi_s \nu_s\, ds + \int_0^{\cdot} (1 - \nu_s \phi_s) d\tilde{N}_s\right\} \prod_{n=1}^{N_t} \phi_{T_n}$$

In order to prove that (Y, Z, U, Q) is a measure solution, we need to show only that $Q = V_T \cdot P$ is a probability measure. From assumption (MS2),

$$\|\frac{f(t, Z_t, U_t)}{Z_t}\| \leq \Phi_t, \quad t \in [0, T]. \tag{4.2.3}$$

Therefore $\int_0^{\cdot} \frac{f(t,Z_t,U_t)}{Z_t}\, dB_t$ is a BMO-martingale with respect to P. By Theorem A.1.2, $\mathcal{E}(M)$ is uniformly integrable. In addition, we can write

$$\int_0^t (1 - \nu_s \phi_s)\, dN_s = \int_0^t \int_{\mathbb{R}^*} (1 - \nu_s \phi_s(q)) p(ds, dq) = \sum_{n=1}^{N_t} \nu_{T_n} \phi(T_n, q_n).$$

Since $(N_t)_{t \in [0,T]}$ is almost surely finite, the above integrals converge if ϕ is bounded on $[0, T]$. In particular, we choose ϕ to be continuous and to satisfy $\phi > 1$. The uniform integrability of V follows from Hölder's inequality, and Q is a probability measure equivalent fo P.

The proof of the converse is the same as in Theorem 4.2.1. □

It is worth noting that unlike the case of BSDEs driven only by Brownian motion, even if ξ is bounded, there are infinitely many measure solutions in the case of BSDEs with jumps. This is due to the fact that the measure Q depends on ϕ.

The next section presents a construction of a measure solution when we do not assume any knowledge of the strong solution.

4.3 Construction of a Measure Solution in the Lipschitz Case

In this section we present a construction of measure solutions that also provides an existence result. This construction was introduced by Ankirchner *et al.* (2009), it consists in iterating a procedure of projection and representation, respectively. The method is particularly interesting in the sense that it is intrinsic, i.e. it does not involve any knowledge of the theory on strong solutions. We consider a one-dimensional Lipschitz BSDE with bounded terminal condition $\xi \in L^2(\Omega)$. We further assume that f is measurable, $f(t,.)$ is almost everywhere continuous and $E\left[\int_0^T |f(s,0)|^2\,ds\right] < \infty$. For simplicity, we put $f(t,0) = 0$ for all t and thus define g by: $\forall (t,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}$,

$$g(t,z) = \begin{cases} \frac{f(t,z)}{z} & \text{if} \quad z \neq 0 \\ 0 & \text{if} \quad z = 0. \end{cases}$$

Since f satisfies the standard Lipschitz condition, g is bounded. Put $Q^0 = P$, the real world probability, $B^0 = B$, the canonical Brownian motion, $Z^0 = Y^0 = 0$ and $\Gamma_t^0 = \int_0^t g(s, Z_s^0)\,dB_s$. Put $Y^1 = E\left[\xi|\mathcal{F}_.\right]$, which is a martingale. By martingale representation, there exists a predictable square integrable process Z^1 such that

$$Y^1 = E\left[\xi|\mathcal{F}_.\right] = E\left[\xi\right] + \int_0^{\cdot} Z_s^1\,dB_s^0.$$

Since g is bounded, Novikov's criterion implies that

$$Q^1 = \mathcal{E}(\Gamma^1)_T \cdot P$$

is a probability, with $\Gamma^1 = \int_0^{\cdot} g(s, Z_s^1)\,dB_s$. By Girsanov's transformation,

$$B^1 = B - \int_0^{\cdot} g(s, Z_s^1)\,ds$$

is a Brownian motion under Q^1. Now, we need to construct a pair of processes (Y^2, Z^2) along with a probability measure Q^2 in the same way that we have done for (Y^1, Z^1) and Q^1. In Brownian filtration frameworks such as ours, every pair of probability measures $Q \sim P$ on $\mathcal{F}_T$ is a Girsanov pair. Hence, (Q^1, Q^0) is a Girsanov pair with random density $\frac{dQ^1}{dQ^0} = \mathcal{E}(\Gamma^1)_T$. From Revuz and Yor (1999), see Exercise 1.27, since the martingale Y^1 admits a representation by a predictable process under Q^0, its analogue $Y^2 = E^{Q^1}\left[\xi|\mathcal{F}_.\right]$ is also representable by a predictable process under the measure Q^1. In other words, there exists a predictable process Z^2 such that $Y^2 = E^{Q^1}\left[\xi|\mathcal{F}_.\right] = E^{Q^1}[\xi] + \int_0^t Z_s^2\,dB_s^1$. By the same process, we recursively define the sequences $(Q^n)_{n\in\mathbb{N}}$, $(B^n)_{n\in\mathbb{N}}$, $(Z^n)_{n\in\mathbb{N}}$ and $(Y^n)_{n\in\mathbb{N}}$. Put for all $n \in \mathbb{N}$ $R_T^n = \mathcal{E}(\Gamma^n)_T$, with $\Gamma^n = \int_0^{\cdot} g(s, Z_s^n)\,dB_s$, the Radon-Nikodym derivative of Q^n by P. The self-contained existence result of measure solutions is thus given as follows:

Theorem 4.3.1. *If f is adapted and Lipschitz, and assumption (MS2) holds, then the sequence $(Y^n, Z^n)_{n\in\mathbb{N}}$ converges in BMO to the pair of processes (Y, Z) such that Z is adapted and square integrable, $Q = \mathcal{E}(\int_0^{\cdot} g(s, Z_s)\,dB_s)_T \cdot P$ is a probability measure equivalent to P and (Y, Z, Q) defines a measure solution of the Lipschitz BSDE with a bounded terminal condition.*

Proof. See Ankirchner *et al.* (2009), Theorem 3.1. □

This self-contained construction of measure solutions of BSDEs with Lipschitz continuous generators and bounded terminal conditions can be combined with an appropriate time discretization to construct a numerical scheme.

4.4 Approximating Measure Solutions

In this section, we shall assess the approximation of measure solutions of Lipschitz BSDEs. Following Section 4.3, the solution (Y, Z, Q) is the limit of a sequence (Y^n, Z^n, Q^n) such that for all $n \in \mathbb{N}$,

$$Y^{n+1} = E^{Q^n}[\xi|\mathcal{F}_\cdot] = E^{Q^n}[\xi] + \int_0^\cdot Z_s^{n+1}\, dB_s^n, \tag{4.4.1}$$

Z^{n+1} defined by martingale representation, and Q^n the equivalent probability measure that makes $B^n = B - \int_0^\cdot g(s, Z_s^n)\, ds$ a Brownian motion. The Radon-Nikodym derivative of Q^n with respect to P is given by $R_T^n = \mathcal{E}(\int_0^T g(s, Z_s^n)\, dB_s)$. Let us now introduce an explicit time discretization of the above iteration.

Consider a partition $p = \{t_0, t_1, \ldots, t_N\}$ of $[0, T]$ with mesh size $|p| = \max_i |t_{i+1} - t_i|$. Define $\Delta_i = t_{i+1} - t_i$ and $\Delta B_i = B_{t_{i+1}} - B_{t_i}$ (we take $B_0 = B_{t_0} = 0$). Let $n \in \mathbb{N}^*$ and $i \in \{0, \ldots, N\}$. Then, $Y_{t_i}^{n+1} = E^{Q^n}[\xi|\mathcal{F}_{t_i}]$. The martingale representation theorem ensures the existence of a process $(\tilde{Z}_t^n)_{t\in[0,T]}$ such that

$$Y_{t_i}^{n+1} = Y_{t_{i+1}}^{n+1} - \int_{t_i}^{t_{i+1}} \tilde{Z}_s^n\, dB_s^n. \tag{4.4.2}$$

The random variable $Z_{t_i}^n$ is taken as the best approximation of $\tilde{Z}_{t_i}^n$ by a $\mathcal{F}_{t_i}$-measurable random variable, i.e.

$$Z_{t_i}^{n+1} = (t_{i+1} - t_i)^{-1} E^{Q^n}\left[Y_{t_{i+1}}^{n+1}(B_{t_{i+1}}^n - B_{t_i}^n)|\mathcal{F}_{t_i}\right].$$

To have the relation above, we multiply both sides in (4.4.2) by $\int_{t_i}^{t_{i+1}} dB_s^n$, and we use Itô isometry along with the fact that $Y_{t_i}^{n+1}$ is $\mathcal{F}_{t_i}$-measurable. Using the first equality in (4.4.1) with the tower property of the conditional expectation, we have the simpler expression: $Z_{t_i}^{n+1} = (t_{i+1} - t_i)^{-1} E^{Q^n}[\xi \Delta B_i^n|\mathcal{F}_{t_i}]$. Thus, we are led to the following numerical scheme:

$$\begin{cases} Y^0 & = Z^0 = 0 \\ Z_{t_i}^{n+1} & = (t_{i+1} - t_i)^{-1} E^{Q^n}\left[\xi \Delta B_i^n \mid \mathcal{F}_{t_i}\right] \\ Y_{t_i}^{n+1} & = E^{Q^n}[\xi] + \sum_{k=0}^i Z_{t_k}^{n+1} \Delta B_k^n \\ R_{t_N}^{n+1} & = \exp\left\{\sum_{i=0}^N g(t_i, Z_{t_i}^n)\Delta B_i - \frac{1}{2}\sum_{i=0}^N g(t_i, Z_{t_i}^n)^2 \Delta_i\right\} \\ Q^{n+1} & = R_{t_N}^{n+1} Q^n \\ B_{t_i}^{n+1} & = B_{t_i}^n - \sum_{k=0}^i g(t_k, Z_{t_k}^n)\Delta_k. \end{cases} \tag{4.4.3}$$

Our numerical scheme cannot be implemented as it stands. The major problem is that the expectations are taken under new probabilities, and it will not be possible to generate random numbers under these probabilities with standard numerical packages. Therefore, we need to express the expectations in terms of expectations under the "real world" probability. In that regard, the formula $E^Q[\xi] = E\left[\xi \frac{dQ}{dP}\right]$ and the following result will be useful:

Proposition 4.4.1. *Let $\mathcal{G} \subset \mathcal{F}$ be a sub σ-algebra, X an $\mathcal{F}$-measurable random variable, Q a measure defined by $dQ = VdP$, and V $\mathcal{F}$-measurable. Then, we have:*

1. $dQ_{|\mathcal{G}} = E[V \mid \mathcal{G}]dP_{|\mathcal{G}}$
2. $E[V \mid \mathcal{G}]\, E^Q[X \mid \mathcal{G}] = E[VX \mid \mathcal{G}]$

Proof. Let $A \in \mathcal{G}$. We have $E[E[V \mid \mathcal{G}]\, 1_A] = E[E[V1_A \mid \mathcal{G}]] = E[V1_A] = Q(A)$. This shows the first claim.

Let $A \in \mathcal{G}$ and put $L = E[V \mid \mathcal{G}]$, which is $\mathcal{G}$-measurable. We have $E^Q[1_A E[VX \mid \mathcal{G}]] = E[L1_A E[VX \mid \mathcal{G}]]$ due to the first claim of the proposition. Using the fact that L is $\mathcal{G}$- measurable and the tower property, we have $E[L1_A E[VX \mid \mathcal{G}]] = E[L1_A VX] = E^Q[L1_A X]$. By the tower property again we conclude that $E^Q[1_A E[VX \mid \mathcal{G}]] = E^Q\left[1_A E^Q[LX \mid \mathcal{G}]\right]$. Hence, the second claim follows from the partial averaging property and the fact that L is $\mathcal{G}$-measurable. $\square$

Thus, by applying the previous proposition, we are able to eliminate the probability Q^n in the computations. We transform the scheme (4.4.3) into the scheme

$$
\begin{cases}
Y^0 & = Z^0 = 0 \\
Z^{n+1}_{t_i} & = (t_{i+1} - t_i)^{-1} \frac{E\left[\xi \Delta B^n_i R^n_{t_N} \mid \mathcal{F}_{t_i}\right]}{E\left[R^n_{t_N} \mid \mathcal{F}_{t_i}\right]} \\
R^{n+1}_{t_N} & = \exp\left\{\sum_{i=0}^N g(t_i, Z^n_{t_i})\Delta B_i - \frac{1}{2}\sum_{i=0}^N g(t_i, Z^n_{t_i})^2 \Delta_i\right\} \\
Y^{n+1}_{t_i} & = E\left[\xi R^n_{t_N}\right] + \sum_{k=0}^i Z^{n+1}_{t_k} \Delta B^n_k \\
Q^{n+1} & = R^{n+1}_{t_N} Q^n \\
B^{n+1}_{t_i} & = B^n_{t_i} - \sum_{k=0}^i g(t_k, Z^n_{t_k})\Delta_k.
\end{cases}
\tag{4.4.4}
$$

Now we have a numerical scheme for an approximation of measure solutions of Lipschitz BSDEs. This scheme stands out by its self-contained property. We should point out that we still need to approximate the conditional expectations. There is a wealth of literature exploring this question, especially in the approximation of BSDEs. See the introduction of Richou (2010) for a survey of some techniques. Moreover, in the present study we will not address the rather important questions of convergence and estimation of the error of approximation. This is left for future research. Note that due to Theorem 4.2.1, this numerical scheme can also be used to approximate (strong) solutions of Lipschitz BSDEs.

Chapter 5

Applications, Numerics and Conclusion

The results we derived and presented in the previous chapters show how BSDEs can be used to solve a problem of optimal cross hedging. The techniques can be applied to derive further results. In addition there are numerous problems in finance that can be modelled similarly to the cross hedging, and solved with the methods of Chapter 2. On the other hand, the numerical approximation of solutions is an issue of great importance as numerical results are the ones used in industry for decision-making.

5.1 Optimal Hedge

The results of the previous chapters can be used to describe the optimal hedge $(\Delta_t)_{t\in[0,T]}$, i.e. the part of the optimal cross hedging strategy due to the random liability F (see Subsection 1.2.1). This requires us to consider the utility maximization problem with and without terminal liability; we will do so using the characterization of Section 2.3.1. Thus, by Proposition 2.3.10, the value function of the problem with liability respectively, without liability is

$$V^F(x) = -\exp(-\eta(x - Y_0)),$$

and

$$V^0(x) = -\exp(-\eta(x - \hat{Y}_0)).$$

The process $(Y_t)_{t\in[0,T]}$ is the first component of the solution of the BSDE

$$Y_t = F + \int_t^T f(s, Z_s)\, ds - \int_t^T Z_s\, dB_s,$$

and the process $(\hat{Y}_t)_{t\in[0,T]}$ is the first component of the solution of the BSDE

$$\hat{Y}_t = \int_t^T f(s, \hat{Z}_s)\, ds - \int_t^T \hat{Z}_s\, dB_s.$$

The function f is given by Equation (2.3.7). Furthermore, by Proposition 2.3.10 and Remark 2.3.5, the pure investment strategy $\hat{\pi}$ and the optimal cross hedging strategy π are such that

$$\hat{\pi}\sigma \in \Pi_C\left(\frac{\theta}{\eta} + \hat{Z}\right) \quad \text{and} \quad \pi\sigma \in \Pi_C\left(\frac{\theta}{\eta} + Z\right).$$

Considering the definition of B given by Equation (1.2.3), the volatility of S can be written as $\Sigma = (\sigma\rho \quad \sigma\sqrt{1-\rho^2})$. Then Z and $\hat{Z}$ are two-dimensional processes (the integral in the BSDE is now with respect to $W = (W^1, W^2)$, see Subsection 1.2.1). For simplicity we assume C to be convex and of the form $C = \{x\Sigma : x \in \mathbb{R}\}$. Therefore, the orthogonal projection of a vector z onto C is given by $\Pi_C(z) = z\Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma$. Hence,

$$\begin{aligned}
\Delta_t &= \pi_t - \hat{\pi}_t \\
&= (Z_t - \hat{Z}_t)\Sigma^*(\Sigma\Sigma^*)^{-1} \\
&= \frac{\rho_t}{\sigma_t}(Z_t^1 - \hat{Z}_t^1) + \frac{\sqrt{1-\rho_t^2}}{\sigma_t}(Z_t^2 - \hat{Z}_t^2).
\end{aligned}$$

Notice that the optimal hedge is a function of the correlation coefficient. If $Z_t - \hat{Z}_t \geq 0$ a.s., then Δ decreases on $[-1, 0]$ and is not strictly decreasing on $[0, 1]$, but decreases very rapidly close to 1. This shows the expected fact that low correlations lead to a high non-hedgeable risk. On the other hand it is worthwhile saying that changing the Brownian motion from B to W does not affect the results of Chapter 2. Since ρ is deterministic, the filtrations generated by the two Brownian motions are the same. The case of random correlation is treated in Frei (2009).

5.2 Indifference Price of a Defaultable Insurance Contract

In this section, we aim at applying the results of the previous chapters to a rather different problem from the one described in Chapter 1. Basically, we would like to describe the price of an insurance contract which may default a finite number of times in the time interval $[0, T]$. Let us start by presenting our model of defaultable claims, that is an extension of the model of Ankirchner *et al.* (2010*a*).

5.2.1 Model

We are still working on the filtered probability space $(\Omega, \mathcal{F}, P)$ defined in Subsection 1.2.1. Let us assume that it is possible to purchase, in an insurance company, a contract which can default at any time between the initial time and the horizon. However, the event of a default does not lead to the end of the contract, nor to the certitude that another default will or will not occur.

The defaultable claim F is given by the sum

$$F = F^0 \mathbf{1}_{S>T} + \sum_{n \geq 1} F^n \mathbf{1}_{T_n \leq T}, \quad S = \min\{T_n, n \geq 1\}.$$

This means that the trader receives the promised pay-off F^0 (which is a $\mathcal{F}_T$-random variable) at time T if no default occurs up to the horizon time, and at each occurrence of the default prior to time T, he receives F^n (if the default happens at time T_n). It is rather natural to assume that at each time t, the trader knows if a default takes place and if any defaults have already taken place. In that regard, an important issue that needs to be addressed is how to model the knowledge of the investor, because his flow of information is no longer the filtration $\mathbb{F}$. We should, in addition, incorporate the knowledge on the occurrence of the defaults. As quite a number of works dealing with credit risk modelling, see for instance Ankirchner *et al.* (2010a), our study will utilize the technique of progressive enlargement of filtration. Let k_T be the total number of observed defaults, and define $\mathbb{G} = (\mathcal{G}_t)_{t \in [0,T]}$ (named the full filtration) by

$$\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\mathbf{1}_{T_1 \leq t}) \vee \cdots \vee \sigma(\mathbf{1}_{T_{k_T} \leq t}), \quad t \in [0, T].$$

This is actually k_T progressive[1] enlargements of the (reference) filtration $\mathbb{F}$. Before going any further in the description of the model, let us point out some noteworthy facts about the full filtration.

Remark 5.2.1.

- *The theory of progressive enlargement of filtration, largely developed by Yor and Jeulin, was introduced to study properties of non-stopping random times. For further reading on the topic, we refer to Jeulin (1979) and Jeulin and Yor (1978), or to Mansuy and Yor (2006) for a text written in English.*

- *The random times $(T_n)_{n \geq 1}$ are not stopping times under the filtration $\mathbb{F}$. However, they are stopping times under $\mathbb{G}$. We can also define $\mathbb{G}$ as the smallest filtration containing $\mathbb{F}$ and that makes all the T_n stopping times. This is explained, in the case of one enlargement by Coculescu* et al. *(2008).*

- *The occurrence of a default is assumed to be known by all the agents acting on the market. Therefore, a trader who uses this information in his investment process should not be able to construct an arbitrage. Thus, it is important to make sure that the enlargement of filtration keeps the market arbitrage-free.*

[1] The progressively enlarged filtration is usually defined as $\mathcal{F}_t \vee \sigma(\mathbf{1}_{\tau \leq t})$, where $\tau : \Omega \to [0, \infty]$ is the random default time, see for instance Bielecki and Jeanblanc (2009).

One usually assumes the so-called immersion property of the enlarged filtration in financial modelling, stated as follows:

(**H**) Any square integrable $(\mathbb{F}, P)$-martingale is a square integrable $(\mathbb{G}, P)$-martingale.

The (**H**) hypothesis is useful to have a martingale representation in the full filtration, and also helps to derive the arbitrage-free property of the market. For a more thorough study of the subject, we refer to Jeanblanc and Cam (2009). The reader may find a detailed study of the financial relevance of the immersion property in the work of Coculescu *et al.* (2008). We do not assume the (**H**) hypothesis, we rather assume the following "density hypothesis":

(**DH**) There exists an $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^+)$-measurable function $(\omega, q) \mapsto \alpha_t(\omega, q)$ such that

$$P\left[(T_1, \ldots, T_n) \in dq | \mathcal{F}_t\right] = \alpha_t(q) n(dq).$$

The investor with initial wealth v has an exponential utility $-\exp(-\eta x)$ $(\eta \in (0,1))$. We assume that he buys the insurance contract for p and invests $x = v - p$ in a jump financial market with a risky asset whose discounted prices $(S_t)_{t\in[0,T]}$ follow the dynamics:

$$dS_t = S_{t-}\left(\mu\, dt + \sigma\, dB_t + dJ_t\right),$$

with $\mu \in \mathbb{R}_+$, $\sigma \in \mathbb{R}_+^*$ bounded, J the jump process defined in Subsection 4.2.1, and a risk-less bond whose interest rate is therefore assumed to be at every time $r = 0$. Delbaen and Schachermayer (1994) proved that the absence of arbitrage is related to the martingale property of the price dynamics of the assets. Using Girsanov's theorem and suitable conditions on the market parameters, there exists $Q \sim P$ such that S is a $\mathbb{G}$-martingale under Q (see Pham (2010) and the references therein). The trader runs an investment strategy $(\pi_t)_{t\in[0,T]}$ which is —we emphasize this point— assumed to be $\mathbb{G}$-predictable, satisfies the usual admissible conditions (see Subsection 1.2.1). We denote the set of admissible strategies $\mathcal{A}_{\mathbb{G}}$. The constrained set is still closed. Put

$$V^F(x) = \sup_{\pi\in\mathcal{A}_{\mathbb{G}}} E\left[-\exp\left\{-\eta(X_T^{x,\pi} + F)\right\}\right]; \tag{5.2.1}$$

where $(X_t^{x,\pi})_{t\in[0,T]}$ is his wealth process, satisfying

$$X_t = x + \int_0^t \pi_s\sigma_s(\theta_s\, ds + dB_s) + \int_0^t \pi_s \int_{\mathbb{R}^*} U_s(q)\tilde{p}(ds, dq).$$

Since the investment strategies are subject to constraints, the market is incomplete. The problem is to find the indifference (buying) price of the defaultable claim F i.e. the real number $p(v)$ satisfying

$$V^0(v) = V^F(v - p(v)). \tag{5.2.2}$$

Under (**H**) hypothesis, this pricing problem is usually solved by expressing the value function of Problem (5.2.1) in terms of the solution of a BSDE with jumps, see Ankirchner *et al.* (2010*a*). Because we would like to deal (in the BSDE characterization) with BSDEs driven by only Brownian motion, we use the results presented in Appendix A Section A.2 to decompose Problem (5.2.1) into sub-problems to be solved in the reference filtration. For exposition purposes, we shall assume that there are $N = 2$ possible defaults.

5.2.2 BSDE Characterization

We start by decomposing the control problem (5.2.1) into sub-problems. From Lemma A.2.1, any $\pi \in \mathcal{A}_{\mathbb{G}}$ can be written as $\pi = (\pi^0, \pi^1, \pi^2) \in \mathcal{A}^0_{\mathbb{F}} \times \mathcal{A}^1_{\mathbb{F}} \times \mathcal{A}^2_{\mathbb{G}}$ (see the comment after Lemma A.2.1), and the wealth process takes the form $X^\pi_t = X^{\pi^0}_t 1_{t \le T_1} + X^{\pi^1}_t 1_{T_1 < t \le T_2} + X^{\pi^2}_t 1_{t \ge T_2}$. From Theorem A.2.2 the value function V^F is obtained in three steps: solve the problem after the two defaults have occurred, solve the problem in between the two defaults and the before-default problem. Hence, we need to solve

$$V^F_2(x, \delta) = \sup_{\pi^2 \in \mathcal{A}^2_{\mathbb{F}}} E\left[- \exp\left\{-\eta \left(X^{\pi^2}_T(\delta) + F^2(\delta)\right)\right\} \alpha_T(\delta) \mid \mathcal{F}_{\delta_2}\right]; \qquad (5.2.3)$$

$$\begin{aligned} V^F_1(x, \delta_1) = \sup_{\pi^1 \in \mathcal{A}^1_{\mathbb{F}}} E\Big[&- \exp\left\{-\eta(X^{\pi^1}_T(\delta_1) + F^1(\delta_1)\right\} \alpha^1_T(\delta_1) \\ &+ \int_{\delta_1}^T V^F_2(X^{\pi^1}_{\delta_2}, \delta_1, \delta_2)\, d\delta_2 \mid \mathcal{F}_{\delta_1}\Big] \end{aligned} \qquad (5.2.4)$$

and

$$V^F_0(x) = \sup_{\pi^0 \in \mathcal{A}^0_{\mathbb{F}}} E\left[- \exp\left\{-\eta \left(X^{\pi^0}_T + F^0\right)\right\} \alpha^0_T + \int_0^T V^F_1(X^{\pi^0}_{\delta_1}, \delta_1)\, d\delta_1\right], \qquad (5.2.5)$$

with $\delta = (\delta_1, \delta_2)$. In the sequel, we assume the positive terms α_T and α^k_T, $k = 1, 2$, to be equal to 1. This choice is still general. Indeed, by transforming the claim into $\tilde{F}^2(\delta) = F^2(\delta) + \frac{1}{\eta}\log(\alpha_T)$, the term α_T is cancelled. We can apply the same reasoning to the terms α^k_T.

1. The after-defaults problem.

   Problem (5.2.3) is the hedging problem studied in Chapter 2. Here we use the method via Itô-Ventzell's formula. Theorem 2.3.14 entails that

   $$V^F_2(x, \delta) = e^{-\eta x} Y^2_{\delta_2},$$

   where Y^2 is the first component of the unique solution of the BSDE

   $$Y^2_t = -e^{\eta F^2(\delta)} - \int_t^T \frac{|\theta_s Y^2_s + Z^2_s|^2}{Y^2_s}\, ds - \int_t^T Z^2_s\, dB_s - M^2_t.$$

2. The in-between-defaults problem.

   We also solve Problem (5.2.4) using the method of Subsection 2.3.2. The value function takes the form

$$V_1^F(x,\delta_1)$$
$$= \sup_{\pi^1\in\mathcal{A}^1_{\mathbb{F}}} E\left[-\exp\left\{-\eta(X_T^{\pi^1}(\delta_1)+F^1(\delta_1))\right\}+\int_{\delta_1}^T e^{-\eta X_{\delta_2}^{\pi^1}} Y_{\delta_1}^2\, d\delta_2 \mid \mathcal{F}_{\delta_1}\right].$$

   Despite the extra term given by the integral of $V_2^F(X_\delta^{\pi^1},\delta)$ in the expression of $V_1^F(x,\delta_1)$, one can still show that $(V_1^F(X_s^{\pi^1},s))_{s\in[\delta_1,T]}$ is a supermartingale using the reasoning of Proposition 2.2.4. Hence, from Theorem 2.3.14

$$V_1^F(x,\delta_1)=e^{-\eta x}Y_{\delta_1}^1,$$

   where Y^1 is the first component of the unique solution of the BSDE

$$Y_t^1=-e^{\eta F^1(\delta_1)}-\int_t^T \frac{|\theta_s Y_s^1+Z_s^1|^2}{Y_s^1}\, ds-\int_t^T Z_s^1\, dB_s-M_t^1.$$

3. The before-defaults problem.

   Note that the solution of the problem before defaults is again the solution of the global problem (5.2.1). We consider the dynamical version of (5.2.5) given by

$$V_0^F(x,t)=\sup_{\pi^0\in\mathcal{A}^0_{\mathbb{F}}} E\left[-\exp\left\{-\eta(X_T^{\pi^0}+F^0)\right\}+\int_t^T e^{-\eta X_{\delta_1}^{\pi^0}} Y_{\delta_1}^1\, d\delta_1 \mid \mathcal{F}_t\right].$$

   This is solved as the case of the in-between-defaults problem. Therefore, $V^F(x)=V_0^F(x)=e^{-\eta x}Y_0^0$ where Y^0 is the first component of the solution of

$$Y_t^0=-e^{\eta F^0}-\int_t^T \frac{|\theta_s Y_s^0+Z_s^0|^2}{Y_s^0}\, ds-\int_t^T Z_s^0\, dB_s-M_t^0.$$

4. The indifference price.

   The final step is to derive the indifference buyer price of the defaultable insurance derivative. From (5.2.2) and the multiplication property of the exponential function, the indifference price p is

$$p=\frac{1}{\eta}\log\left(\frac{V^F(0)}{V^0(0)}\right).$$

   Since the utility function is increasing, $V^F\geq V^0$, i.e. $p\geq 0$. By Theorem 2.3.14, $V^0=e^{-\eta x}Y_0$, with

$$Y_t=-1-\int_t^T \frac{|\theta_s Y_s+Z_s|^2}{Y_s}\, ds-\int_t^T Z_s\, dB_s-M_t.$$

Hence, $p = \frac{1}{\eta}\log(\frac{Y_0^0}{Y_0})$. Note that the processes M and M^k, $k = 0, 1, 2$, in the BSDEs are different one from another as a consequence of the uniqueness of Doob-Meyer decomposition (2.3.18) of a supermartingale.

5.3 Numerics for Quadratic BSDEs

In most of the cases it is hopeless to have a closed form formula of the solution of a quadratic BSDEs. Therefore it is important to construct fast[2] and accurate numerical schemes to approximate quadratic BSDEs, especially in financial applications as the solutions enable us to make decisions. The aim of this section is to present some methods of approximation of quadratic BSDEs and to actually implement them.

5.3.1 Known Results

The approximation of BSDEs is an active and fertile area of research. Nonetheless in nearly all the schemes available, the Lipschitz generator hypothesis is needed to prove convergence or error estimate. Since there are very few methods of approximation of quadratic BSDEs, we will quote them all. Let us consider the FBSDE (3.1.4)-(3.1.5) with f quadratic and g bounded, Lipschitz and α-Hölder. $0 < \alpha \leq 1$.

Richou (2010) goes from the approximating BSDE

$$Y_t^N = g_N(X_T) + \int_t^T f(s, X_s, Y_s^N, Z_s^N)\, ds - \int_t^T Z_s^N\, dB_s,$$

with g_N a Lipschitz approximation of g to define a new time discretization scheme. The main idea is to provide an estimate of Z that depends only on time; this implies to considering that for $t < T$, $f(t, \cdot, \cdot, \cdot)$ is Lipschitz with respect to z with the Lipschitz constant a function of time. Therefore the BSDE to be approximated becomes

$$Y_t^{N,\epsilon} = g_N(X_T) + \int_t^T f^\epsilon(s, X_s, Y_s^{N,\epsilon}, Z_s^{N,\epsilon})\, ds - \int_t^T Z_s^{N,\epsilon}\, dB_s$$

with generator $f^\epsilon(s, x, y, z) = \mathbf{1}_{s<T-\epsilon} f(s, x, y, z) + \mathbf{1}_{s \geq T-\epsilon} f(s, x, y, 0)$, and $\epsilon \in (0, T)$. The modified FBSDE is approximated using a time grid with $2n + 1$ points and with more points close to T, defined as follows:

$$\begin{cases} t_k &= T(1 - \left(\frac{\epsilon}{T}\right)^{k/n}), \quad 0 \leq k \leq n \\ t_k &= T - \left(\frac{2n-k}{n}\right)\epsilon, \quad n \leq k \leq 2n. \end{cases}$$

The approximation $(Y^{N,\epsilon,n}, Z^{N,\epsilon,n})$ of $(Y^{N,\epsilon}, Z^{N,\epsilon})$ is obtained by a slight change of the usual dynamic programming equation (see Richou (2010) Equations

[2] In terms of convergence and running time.

(5.14) page 85 and (1.6) page 5). Let $e(n)$ be the global error estimate. For all η, there exists C such that $e(n) \leq C(n^{-(\alpha-\eta)})$, see Richou (2010) Theorem 5.23.

The truncation procedure is a method introduced by Imkeller and Reis (2010) that consists in approaching the solution of the FBSDE with the solution (Y^n, Z^n) of the FBSDE with the backward part

$$Y_t^n = g(X_T) + \int_t^T f_n(s, X_s, Y_s^n, Z_s^n)\, ds - \int_t^T Z_s^n\, dB_s, \tag{5.3.1}$$

with $f_n(\cdot,\cdot,\cdot,\cdot) = f(\cdot,\cdot,\cdot,h_n(\cdot))$ and $(h_n)_n$ defined in the proof of Theorem 3.4.4. The composition of the functions $f(t,x,y,\cdot)$ and h_n is called the truncation procedure, as it truncates the quadratic growth part of the driver and makes f_n Lipschitz. Moreover, $(f_n)_n$ converges to f (see the proof of Theorem 3.4.4). Imkeller and Reis (2010) show that the approximating error satisfies: for all $p \geq 2$, there exists $C > 0$ such that $e(n) \leq Cn^{-12}$. However, none of the works presenting this method (see Imkeller and Reis (2010), Imkeller *et al.* (2010), Reis (2010)) actually implements it with an example as in the next subsection.

When the generator of the BSDE is of the form

$$f(t,x,y,z) = l(t,x,y) + a(t,z) + \frac{\gamma}{2}|z|,$$

where $\gamma \neq 0$, l and a are measurable functions such that l is Lipschitz in x and y and a is Lipschitz homogeneous in z, it is possible to transform the quadratic BSDE into a BSDE with linear growth in Z. This is done by the exponential (or Cole-Hopf) transformation $P = \exp(\gamma Y)$ and $Q = \gamma P Z$. Hence the BSDE becomes

$$P_t = e^{\gamma g(X_T)} + \int_t^T \left[\gamma P_s f\left(s, X_s, \frac{log(p_s)}{\gamma}, \frac{Q_s}{\gamma P_s}\right) - \frac{1}{2}\frac{|Q_s|^2}{P_s}\right] ds - \int_t^T Q_s\, dB_s.$$

The solution (Y, Z) is retrieved by applying the inverse transformation. Imkeller *et al.* (2010) obtained the error estimate $e(n) \leq C(\pi + \pi^{1-\epsilon})$, $\epsilon > 0$ and π the mesh size of the partition of $[0, T]$.

The reader will notice that all the techniques described above strive to turn the problem of approximating a quadratic BSDE into the approximation of a Lipschitz BSDE. This is because the theory of numerics for BSDE with linear growth is rather well established. We refer to the introduction of Richou (2010) for a presentation of the principal schemes for Lipschitz BSDEs.

5.3.2 Application to a Pricing Problem

We will apply the Cole-Hopf transformation and the truncation procedure to solve the problem of pricing of an insurance derivative using an exponential utility maximization. First of all, let us mention that in the subsequent implementations, we use the algorithm by Crisan and Manolarakis (2010) to approximate the BSDEs with linear growth in Z. This is done with a slight difference: Instead of approximating the expectations by the Cubature method as described in the paper, we use the Ninomiya-Victoir scheme (see Matchie (2010) for a presentation of both of the schemes). We test this implementation on the valuation of a call option in a Black-Scholes model as studied in Subsection 3.1.1, but with no consumption process. This corresponds to solving the linear BSDE (3.1.2) with terminal condition $V_T = (X_T - K)^+$ where K is the strike and X_T the terminal value of the wealth process. The process $(X_t)_{t\in[0,T]}$ is defined[3] by

$$dX_t = \mu X_t\,dt + \sigma X_t\,dB_t = (\mu - \frac{\sigma^2}{2})X_t\,dt + \sigma X_t \circ dB_t. \tag{5.3.2}$$

Since we know a closed form formula for the solution V_0, given by Equation (3.1.3), we can compute the approximating error depicted by Figure 5.1. The expectation of (3.1.3) is approximated by the Euler-Maruyama scheme.

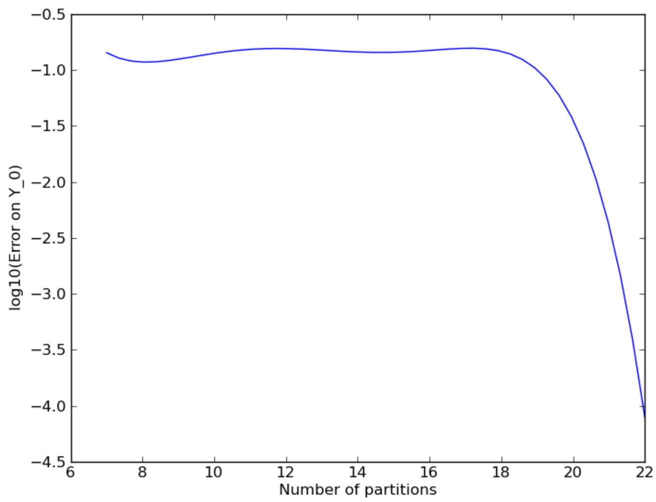


Figure 5.1: Approximating error with the market parameters: $\mu = 0.05$, $r = 0.07$, $\sigma = 0.2$ and $X_0 = K = 10$.

Compared to the numerical results of Crisan and Manolarakis (2010), we are able to achieve a similar order of approximating error for a higher number of partitions. However, this remains to be proved theoretically.

[3]The sign "$\circ$" denotes the stochastic integral in the Stratonovich form.

Remark 5.3.1 (About the implementation)**.** *We do not have to approximate the solutions of the ODEs needed in the implementation of the Ninomiya-Victoir scheme as, in our case, all of the three ODEs have a closed form solution. To ease the implementation we notice that we have to transform the process* (X_t) *into a process driven by a 2-dimensional Brownian motion. More precisely, we write*

$$dX_t = \mu X_t\,dt + \sigma\alpha X_t\,dB_t^1 + \sigma X_t\sqrt{1-\alpha^2}\,dB_t^2, \quad |\alpha| \le 1,$$

which is equivalent to Equation (5.3.2) *by the Lévy characterization of Brownian motion (*B^1 *and* B^2 *being independent). We simulate the algorithm with* 30000 *sample paths and we launch the simulation 10 times. The graph is obtained after a polynomial interpolation that helps to smooth out the curve.*

In the rest of this subsection we apply the Cole-Hopf transformation and the truncation procedure to numerically price an European put-option on kerosene, in the light of Imkeller *et al.* (2010). Note that kerosene is not traded in a liquid market, and its price is known to be highly correlated with the price of heating oil. The price of kerosene and heating oil, respectively, follow the dynamics:

$$dR_t = \mu(t, R_t)\,dt + \sigma(t, R_t)\,dB_t^1 = 0.12R_t\,dt + 0.41R_t\,dB_t^1$$

$$dS_t = \alpha S_t\,dt + \beta S_t\,dB_t = 0.1\,dt + 0.35\,dB_t,$$

with the spot prices r_0 and $s_0 = 173$ money units. The price of the contract is given by $p_t = \frac{1}{\eta}\log\left(\frac{V_t^0}{V_t^F}\right) = Y_t^F - Y_t^0$ where Y^F is the first component of the solution of the BSDE

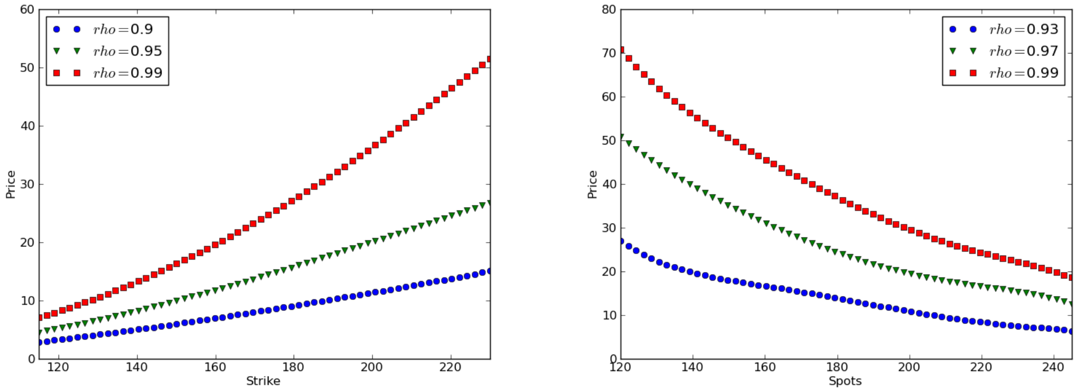
$$Y_t = F(R_T) + \int_t^T f(s, R_s, Z_s)\,ds - \int_t^T Z_s\,dB_s^1, \quad t \in [0, T].$$

With $f(t, r, z) = \frac{\theta(t,r)}{2\eta} - z\rho\theta(t, r) - \frac{\eta}{2}(1-\rho^2)z^2$ and $F(r) = (K - r)^+$ for a given strike price K, and Y^0 solution of the BSDE with same generator and terminal condition $\xi = 0$, see Imkeller *et al.* (2010), Lemma 1.

The quadratic BSDE is first solved by the Cole-Hopf transformation, which leads to the BSDE

$$P_t = \exp\{cF(R_T)\} + \int_t^T \left(\frac{\theta^2}{2\eta}cP_s - \rho\theta Z_s\right)\,ds - \int_t^T Q_s\,dB_s, \tag{5.3.3}$$

$c = -\eta(1-\rho^2)$. This last BSDE is solved using the Crisan-Manolarakis scheme and the solution of the actual BSDE is obtained by the inverse transformation. We launch the algorithm for different values of ρ and plot the time 0 price for varying strikes and kerosene spot levels.



(a) Value of the put-option in terms of the strike price at the spot $r_0 = 170$.

(b) Value of the put-option in terms of the spot price at the strike $K = 200$.

Figure 5.2: Price of the put option for varying kerosene spot and strike and at different levels of correlation.

The graphs reveal that low correlation levels lead to lower prices of the contract. As explained by Imkeller *et al.* (2010), this is because when the correlation between the non-tradable asset and the tradable one decreases, the non-hedgeable residual risk increases (see figure 1.1.2) and therefore leads to low prices (because the risk-covering effect is low). This is better explained by the price surface below.

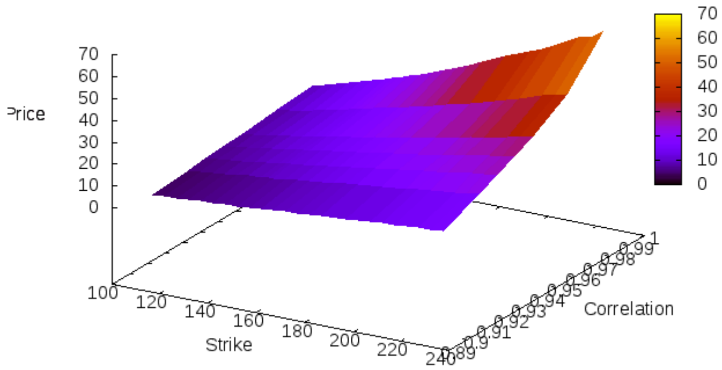


Figure 5.3: Value of the put-option at time $t = 0$ for varying strike price and correlation level for the fixed spot $r_0 = 170$.

We implement, in addition, the truncation procedure to find the price p_0 of the put-option on kerosene. The truncating function is defined by

$$h_n(z) = \begin{cases} n+1 & z > n+2 \\ z & |z| \leq n \\ -(n+1) & z < -(n+2) \\ (-n^2 + 2nz - z(z-4))/4 & z \in [n, n+2] \\ (n^2 + 2nz + z(z+4))/4 & z \in [-(n+2), -n]. \end{cases}$$

We solve the BSDE (5.3.1) for each n and stop the iterations when the distance between two subsequent time-zero values is less than 0.00915. The same lines of code are used to solve both Equation (5.3.3) and Equation (5.3.1) for each n.

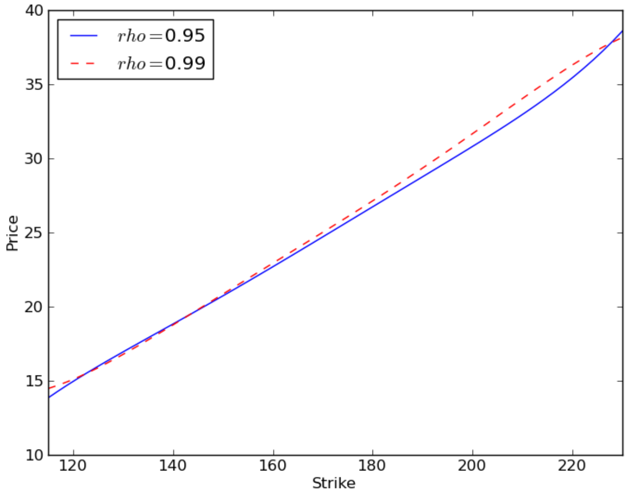


Figure 5.4: Value of the put-option at time $t = 0$ (using the truncation procedure) for varying strike price and different level of correlation for the fixed kerosene spot $r_0 = 170$.

The algorithm of the truncation procedure tends to be very sensible to the partition of the time interval. Here we use the partition suggested in the scheme by Richou (2010) (see the previous subsection), this leads to a result a bit different from the result by the equidistant partition. Moreover, the convergence of the scheme is rather slow. Around 50 iterations are necessary for the converge, therefore the computational cost is relatively high compared to the Cole-Hopf transformation. Furthermore, the scheme seems not to be very sensible to the change of correlation.

For the sake of comparison, we plot on the same graph the price of the put option for varying spot prices using the truncation procedure and the Cole-Hopf transformation.

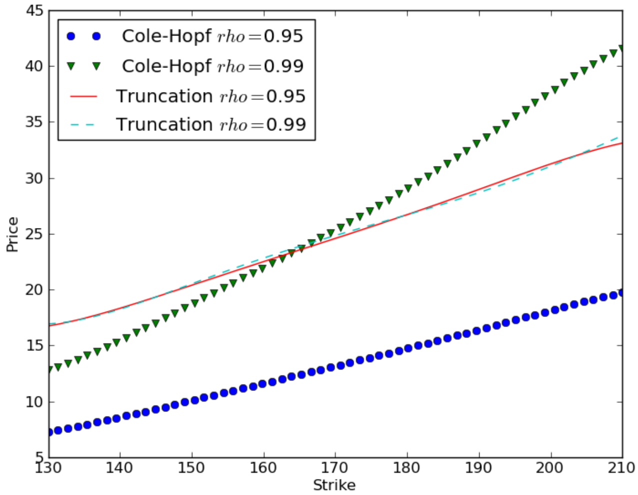


Figure 5.5: Value of the put-option at time $t = 0$ for varying strike price and different level of correlation for the fixed kerosene spot $r_0 = 170$.

5.4 Conclusion

The main concerns of this work have been to present some fully probabilistic methods to hedge non-tradable risks, and to study BSDEs, especially those that have a quadratic generator. To stress the importance of quadratic BSDEs, we have also investigated a wide range of applications of the theory, such as non-linear PDEs, behavioural finance and pricing of defaultable claims in a jump market. The principal results of the thesis span over three interconnected areas: Finance, Stochastic Analysis and Numerical Analysis.

In Finance, the results of the numerical simulations show that the price of the insurance derivative is an increasing function of the correlation between the hedged asset and the hedging instrument used to cross-hedge the contract. In Stochastic Analysis, the results of Chapters 2, 3 and 4 show that the martingale optimality principle, the martingale representation theorem and Itô-Ventzell's formula can be used solely together with Itô's formula and some arguments of Functional Analysis to transform a stochastic control problem into a quadratic BSDE. On the other hand, for BSDEs with bounded terminal conditions, there is an equivalence between the existence of classical solutions and measure solutions of BSDEs, even for BSDEs driven by a compound Poisson process. In

Numerical Analysis, Section 4.4 shows that the self-contained construction of measure solutions can be used to derive an implementable numerical scheme.

From this point, there is a wealth of questions that can be investigated and that we have not looked at. We have studied quadratic BSDEs driven by Brownian motion; it will be interesting to study more general classes of BSDEs, like semimartingale quadratic BSDEs and stochastic partial differential equations driven by Lévy dynamics. In Chapter 2 we obtained three equations characterizing the indifference pricing problem. A step further would be to compare the prices obtained by those different equations, using for instance a numerical approximation. In Chapter 4 we have not considered the case where we start with a BSDE determined by a terminal condition and a probability measure equivalent to the "real world" measure and work out the measure solutions. It will also be of interest to study measure solutions under enlarged or shrunk filtrations, and to complete the study of the numerical scheme of Section 4.4 by the estimation of the global error and the convergence rate of the scheme.

Appendices

Appendix A

Some Results from Stochastic Analysis

A.1 Martingales of Bounded Mean Oscillation

The concept of bounded mean oscillation (BMO) was introduced by John and Nirenberg in the early 1960s in classical analysis. It was brought to probability setting by Getoor and Sharpe, in 1972, who proved a duality result between the Hardy space and the space of BMO-functions. BMO-martingales are important in the BSDE characterizations we present in Chapter 2 and are also used to derive the inequality estimates of Chapter 3. We present here some results used in the thesis. They are taken from the lecture notes of Kazamaki (1994).

Definition A.1.1. *Let M be a martingale on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ with $M_0 = 0$, and $1 \le p < \infty$. Let*

$$\|M\|_{BMO_p} = \sup_{\{\tau, \textit{stopping times}\}} \|E\left[|M_T - M_{\tau-}|^p \mid \mathcal{F}_\tau\right]^{1/p}\|_\infty.$$

A BMO_p-martingale is a uniformly integrable martingale that is an element of the class $\left\{M : \|M\|_{BMO_p} < \infty\right\}$.

It is shown, see Kazamaki (1994), that for $p \neq q$ $BMO_p = BMO_q$. Thus, we drop out the subscript and simply write BMO.

Theorem A.1.2. *Let $M \in BMO(P)$, the stochastic exponential $\mathcal{E}(M)$ is uniformly integrable.*

Proof. See Kazamaki (1994), Theorem 2.3. □

Note that the space BMO actually depends on the underlying probability measure. The next proposition provides a result of transformation of BMO by a change of law.

Proposition A.1.3. *Let $M \in BMO(P)$ and Q the probability measure defined by $dQ = \mathcal{E}(M)_T dP$. The map $\Phi : X \mapsto [X, M] - X$ is an isomorphism of $BMO(P)$ onto $BMO(Q)$.*

Proof. See Kazamaki (1994), Theorem 3.6. □

In particular, $[M] - M$ is a $BMO(Q)$-martingale for every $BMO(P)$-martingale M.

In the next section, we introduce a crucial tool for the study of Chapter 5 Section 5.2, the concept of $\mathbb{F}$-decomposition of a $\mathbb{G}$-predictable stochastic process (where $\mathbb{F}$ is the reference filtration and $\mathbb{G}$ the full filtration).

A.2 $\mathbb{F}$-decomposition

The importance of the $\mathbb{F}$-decomposition lies in the fact that the contingent claim F (Section 5.2) is subject to possible defaults. The method we discuss in Chapter 5 consists of solving the problem in between the occurrence of two successive defaults, and using the intermediate solutions to construct the solution of the initial problem. The $\mathbb{F}$-decomposition thus enabled us to decompose Problem (5.2.1), which was to be solved in the full filtration, into classical subproblems that were solved in the reference filtration. The results of this section are taken from Pham (2010), we refer to that paper for a more detailed study.

Lemma A.2.1. *Any $\mathbb{G}$-predictable process Y can be identified to an $(N+1)$-tuple $(Y^0, \dots, Y^N)$ where each Y^k, $k = 0, \dots, N$ is $\mathbb{F}$-predictable, and Y admits the representation*

$$Y_t = Y_t^0 \mathbf{1}_{t \le T_1} + \sum_{k=1}^{N-1} Y_t^k(T_1, \dots, T_k) 1_{T_k \le t \le T_{k+1}} + Y_t^N(T_1, \dots, T_N) 1_{T_N \le t}.$$

Proof. See Pham (2010), Lemma 2.1. □

This lemma provides the most important argument for the decomposition of a $\mathbb{G}$-predictable process. In particular, for any investment strategy $\pi \in \mathcal{A}_{\mathbb{G}}$ there exists $(\pi^0, \dots, \pi^N)$ such that for each $k = 0, \dots, N$, π^k is $\mathbb{F}$-predictable, we put $\pi^k \in \mathcal{A}_{\mathbb{F}}^k$. Therefore, the admissible set is written as a Cartesian product, i.e. $\mathcal{A}_{\mathbb{G}} = \mathcal{A}_{\mathbb{F}}^0 \times \dots \times \mathcal{A}_{\mathbb{F}}^N$, and the set of constraints $C = C^0 \times \dots \times C^N$ with each C^k closed. Before giving the next result which provides the decomposition of the stochastic control problem in the full filtration $\mathbb{G}$ into sub-problems in the reference filtration, let us introduce a new notation. Put

$$\begin{aligned} \alpha_t^0 &= P(T_1 > t \mid \mathcal{F}_t) \\ &= \int_t^\infty \int_{\delta_1}^\infty \dots \int_{\delta_{N-1}}^\infty \alpha_t(\delta_1, \dots, \delta_N)\, d\delta_N \dots d\delta_1, \end{aligned}$$

and for $k = 1, \dots, N-1$

$$\alpha_t^k(\delta_1, \dots, \delta_k) = \int_t^\infty \int_{\delta_{k+1}}^\infty \dots \int_{\delta_{N-1}}^\infty \alpha_t(\delta_1, \dots, \delta_N)\, d\delta_N \dots d\delta_{k+1}.$$

Theorem A.2.2. *The value function V^F is obtained from the backward induction formulae*

$$\begin{aligned} V_N^F(x, \delta) &= \sup_{\pi^N \in \mathcal{A}_{\mathbb{F}}^N} E\left[U(X_T^N(\delta) - F^N(\delta))\alpha_T(\delta) \;\middle|\; \mathcal{F}_{\delta_N}\right] \\ V_k^F(x, \delta^{(k)}) &= \sup_{\pi^k \in \mathcal{A}_{\mathbb{F}}^k} E\Big[U(X_T^k(\delta^{(k)}) - F^k(\delta^{(k)}))\alpha_T^k(\delta^{(k)}) \\ &\qquad + \int_{\delta_k}^T V_{k+1}(\Gamma_{\delta_{k+1}}^{k+1}(X_{\delta_{k+1}}^k), \delta^{(k)})\, d\delta_{k+1} \;\Big|\; \mathcal{F}_{\delta_k}\Big], \end{aligned}$$

with $\delta = (\delta_1, \dots, \delta_N) \in [0, T]^N$ and ordered, and $\delta^{(k)} = (\delta_1, \dots, \delta_k)$. For the sake of notational simplicity, we omit the dependence of X^k on π^k and x.

Proof. See Pham (2010), Theorem 4.1. □

Appendix B

Codes for the Numerical Implementations

In this final appendix, we provide the numerical codes we use to obtain the results of Chapter 5. They are written in the programming language Python.

B.1 Crisan-Manolarakis Scheme

The following code solves a Lipschitz BSDE using the Algorithm by Crisan and Manolarakis where the expectations are approximated using the Ninomiya-Victor scheme.

```
from __future__ import division
from scipy import*
from scipy.stats import *
from random import*
import pylab as py
from scipy.optimize import fixed_point


#Constants of the model
n = 9
T = 1
k = 50000  #Number of sample paths
d = 2
S0 = 10
K = 10
a1 = 0.05
b1 = 0.2
a2 = 0.05
b2 = 0.2
r = 0.07
```

```
sig = 0.2
theta = (a2-r)/b2
K = 10
l = 1
alpha = a2 - (1/2)*(b2**2)
#=====================================================================
#Definition of useful functions.
def E0(s, l):
  #Value of exp(sV0) with initial condition l.
  return l*exp( alpha*s )
def E1(s, l):
  #Value of exp(sV1) with initial condition l.
  return l*exp( sig*b2*s )
def E2(s, l):
  #Value of exp(sV1) with initial condition l.
  c = sqrt( 1 - sig**2 )*b2
  return l*exp( c*s )
#The functions above come from the resolution of the ODEs,
#see Matchie (2010).
def phi(x):
  #Terminal condition.
  return max(0, x - K)
def f(t,x,y,z):
  return -r*y-theta*sig*z
#=====================================================================
#Exact  solution
exact_Y0 = 0
X0 = S0
S = 0
N = 100
h = T/N
Gam0 = 1
Gamma = 0
X_T = 0
exact_Exp = 0
B11 = 0
B22 = 0
for p in range(5):
  S = 0
  for i in xrange(k):
    S2 = 0
    B11 = 0
    B22 = 0
    for l in xrange(N):
```

```
            B11 = B11 + norm.rvs(scale = h)  #An increment of Brownianmotion.
            B22 = B22 + norm.rvs(scale = h)
        Gamma = Gam0*exp( (r- (1/2)*((theta*sig)**2))*T + b2*(sig*B11) )
        #X_T has been put in the Stratonovish form
        X_T = X0*exp( (a2-(b2**2))*T + b2*(sig*B11 + sqrt(1-sig**2)*B22) )
        S = S+ Gamma*phi(X_T)
    ludo = S/k
    exact_Exp = ludo + exact_Exp
exact_Y0 = exact_Exp/5
print exact_Y0
#==========================================================================
#Implementation of the Scheme.

part = range(7,19)
KK = array(part)
Error = zeros(12)
log_Error = zeros(12)
n_err= 0
while n_err < 100:  #100 launches of the algorithm
    Y_0 = zeros(12)
    Y_1 = zeros(12)
    q = 0
    for n in range(7,19):
        X0 = S0
        P = Expect(n-1, X0)[1] #The function Expect is defined in the
        #code of the Cole-Hopf transformation
        R_B = Expect(n-1, X0)[2]
        SumRB = 0
        t = linspace(0,T,n)
        h = T/h
        for i in range(1,n)[::-1]: #To go backward in time
            E = Expect(i, X0)[0]
            X0 = E
            def l(x):
                return P + h*f(t[i], X0, x, (1/h)*R_B) - x
            P = fixed_point(l,0) #gives the fix point of the function
            #l using Picard iterations. This is to get the operator "R"
            # in the scheme.
            for j in xrange(k):
                Delta_B = norm.rvs(scale = (0.25)**2*h)
                SumRB = SumRB + P*Delta_B
            R_B = SumRB/k

        Y_0[q] = log(abs(P - exact_Y0))/log(10)
```

```
        Y_1[q] = abs(P-exact_Y0 )
        q = q + 1
    n_err = n_err + 1
    Error = Error + Y_1
    log_Error = log_Error + Y_0

Error = Error/100
log_Error = log_Error/100
#Print results
```

B.2 Cole-Hopf Transformation

The following code computes de time zero value of the (first component) of the solution of a quadratic BSDE using the exponential change of variable, or Cole Hopf transformation. The code uses the code of the Algorithm by Crisan and Manolarakis to solve the obtained Lipschitz BSDE.

```
#Constants of the model
n = 12
T = 1
k = 60000
d = 2
S0 = 173
r0 = 170  #Kerosene spot price (external risk)
r_mu = 0.12 #Drift of the external risk process
r_sig = 0.41  #Volatility of the external risk process
K = 180       #Strike price
mu = 0.1
sig = 0.35
eta = 0.3 #Risk aversion level
rho = 0.95  #Correlation coefficient
sig = 0.41  #Its value does not mater
theta = (mu)/sig
alpha = r_mu - (1/2)*(r_sig**2)    #a2 - (1/2)*(b2**2)
a = (theta**2)/(2*eta)
b = -rho*theta
c = - eta*(1 - rho**2)
#==========================================================================
#Definition of useful functions.
def phi(x,K):
    #Terminal condition
    c = - eta*(1 - rho**2)
```

```
    return  exp( c*max(0, K - x) )
def psi(x):
    return 1
def f(t,x,y,z):
    '''
    Generator of the BSDE obtained after the exponential change
    '''
    a = theta**2/(2*eta)
    b = -rho*theta
    c = - eta*(1 - rho**2)
    return  a*c*y + b*z
#==============================================================
#Implementation of the Scheme.
def Expect(m, X0,K):
    '''
    The function returns some expectations with X starting at
     time t[m] with value X0. The function returns a list of three
      elements: the first component is E[X_{t_n}^{t_n+1}], the second
      component is E[phi(X_{t_n}^{t_n+1})] and the last component is
     E[phi(X_{t_n}^{t_n+1})*DeltaB_i].
    '''
    steps = 20
    h = T/(n*steps)
    Sum0 = 0
    Sum1 = 0
    Sum2 = 0
    Nino_X = 0
    Nino_B_X = 0
    l3 = zeros(5)
    for p in range(5):
        Nino_X = 0
        Nino_phi_X = 0
        Nino_B_X = 0
        for j in xrange(0,k):
            X = X0
            for i in range(m,m+steps):
                ber = randint(0,1) #0 or 1, Bernoulli distribution
                for l in range(0,d):
                    nor = zeros(d)
                    nor[l] = gauss(0,1) #chooses a normal random variable.
                if ber == 0:
                    X = E0(1/(2*n),E2(nor[1]/sqrt(n),
                    E1(nor[0]/sqrt(n),E0(1/(2*n),X))))
                else:
```

```
                X = E0(1/(2*n),E1(nor[0]/sqrt(n),
                E2(nor[1]/sqrt(n),E0(1/(2*n),X))))
            Nino_X = Nino_X + X
            Nino_phi_X = Nino_phi_X + phi(X,K)
            Delta_B = norm.rvs(scale = h)
            Nino_B_X = Nino_B_X + phi(X,K)*Delta_B
        Res0 = Nino_X/k
        Res1 = Nino_phi_X/k
        Res2 = Nino_B_X/k
        Sum0 = Sum0 + Res0
        Sum1 = Sum1+Res1
        Sum2 = Sum2 + Res2
    approx_Exp = Sum0/5
    approx_phi_Exp = Sum1/5
    approx_Bro_Exp =  Sum2/5
    return [approx_Exp, approx_phi_Exp, approx_Bro_Exp]
X0 = r0
P = Expect(n-1, X0,K)[1]
R_B = Expect(n-1, X0,K)[2]
SumRB = 0
t = linspace(0,T,n)
h = T/(n*50)
for i in range(1,n)[::-1]:
    E = Expect(i, X0,K)[0]
    X0 = E
    def l(x):
        return P + h*f(t[i], X0, x, (1/h)*R_B) - x
    P = fixed_point(l,0)
    for j in xrange(k):
        Delta_B = norm.rvs(scale = h)
        SumRB = SumRB + P*Delta_B
    R_B = SumRB/k
P_0 = P
Y_0 = (1/c)*log(P_0)
print 'The time',0,'value of the process Y is:', Y_0
```

B.3 Truncation Procedure

```
Y_Fpr = 0
epsilon = 0.02
N = 50       #Starting of the iterations
X0 = r0
```

```
Q = Expect(n-1, X0,K, rho)[1]
R_B = Expect(n-1, X0, K, rho)[2]
SumRB = 0
#Solve the BSDE at the first iteration (initial solution)
for i in range(1,n)[::-1]:
  E = Expect(i, X0, K, rho)[0]
  X0 = E
  def l(x):
    return Q + (t[i]-t[i-1])*gen(rho, t[i], X0,
    x, trunc(N,(1/(t[i]-t[i-1]))*R_B)) - x
  Q = fixed_point(l,4)
  for j in xrange(k):
    Delta_B = norm.rvs(scale = (0.25)**2*(t[i]-t[i-1]))
    SumRB = SumRB + Q*Delta_B
  R_B = SumRB/k
Y_Fcur = Q   #Solution of the BSDE at the first iteration
print 'first value', Y_Fcur
while abs(Y_Fcur - Y_Fpr) > epsilon:  #Stop the iteration
#when two subsequent values are less that epsilon
  N = N+1
  #Solve the BSDE at the iteration N+1
  X0 = r0
  Q = Expect(n-1, X0, K, rho)[1]
  R_B = Expect(n-1, X0, K, rho)[2]
  SumRB = 0
  for i in range(1,n)[::-1]:
    E = Expect(i, X0, K, rho)[0]
    X0 = E
    def l(x):
      return Q + (t[i]-t[i-1])*gen(rho, t[i],
      X0, x, trunc(N,(1/(t[i]-t[i-1]))*R_B) ) - x
    Q = fixed_point(l,0)
    for j in xrange(k):
      Delta_B = norm.rvs(scale = (0.25)**2*(t[i]-t[i-1]))
      SumRB = SumRB + Q*Delta_B
    R_B = SumRB/k
  Y_Fpr = Y_Fcur
  Y_Fcur = Q   #Solution of the BSDE at the nth iteration
  print 'value at', N, 'is', Y_Fcur
YF = Y_Fcur
print 'Time-zero value of the first component of the solution:', YF
```

List of References

Ankirchner, S., Blanchet-Scalliet, C. and Eyraud-Loisel, A. (2010 Nov$a$). Credit Risk Premia and Quadratic BSDEs with a Single Jump. *Int. J. Theor. Appl. Finance*, vol. 13, no. 07, pp. 1103 – 1129.

Ankirchner, S. and Imkeller, P. (2011). Hedging with Residual Risk: a BSDE approach. In: Dalang, R., Dozzi, M. and Russo, F. (eds.), *Seminar on Stochastic Analysis, Radom Fields and Applications VI*, pp. 311–326. Springer Bassel AG.

Ankirchner, S., Imkeller, P. and Popier, A. (2008). Optimal Cross Hedging of Insurance Derivatives. *Stoch. Anal. Appl.*, vol. 26, no. 4, pp. 679–709.

Ankirchner, S., Imkeller, P. and Popier, A. (2009). On Measure Solutions of Backward Stochastic Differential Equations. *Stochastic Process. Appl.*, vol. 119, no. 9, pp. 2744–2772.

Ankirchner, S., Imkeller, P. and Reis, G.D. (2010 March$b$). Pricing and Hedging of Derivatives Based on Non-tradable Underlyings. *Math. Finance*, vol. 20, no. 2, pp. 289–312.

Becherer, D. (2006). Bounded Solutions to Backward SDEs with Jumps for Utility Optimization and Indifference Hedging. *Ann. Appl. Probab.*, vol. 16, no. 4, pp. 2027–2054.

Bensoussan, A. (1983). Stochastic Maximum Principle for Distributed Parameter Systems. *Journal of The Franklin Institute*, vol. 315, no. 5-6, pp. 387 – 406. ISSN 0016-0032.

Bielecki, T.R. and Jeanblanc, M. (2009). Indifference Pricing of Defaultable Claims. In: Carmona, R. (ed.), *Indifference Pricing*, pp. 211–230. Princeton University Press.

Bismut, J.-M. (1973). Conjugate Convex Functions in Optimal Stochastic Control. *J. Math. Anal. Appl.*, vol. 44, pp. 384–404.

Briand, P. and Hu, Y. (2006). BSDE with Quadratic Growth and Unbounded Terminal Value. *Probab. Theory Related Fields*, vol. 136, no. 4, pp. 604–618.

Briand, P. and Hu, Y. (2008). Quadratic BSDEs with Convex Generators and Unbounded Terminal Conditions. *Probab. Theory Related Fields*, vol. 141, no. 3-4, pp. 543–567.

Cheridito, P. and Hu, Y. (2010 Oct). Optimal Consumption and Investment in Incomplete Markets with General Constraints. Available at: arXiv:1010.0080v2.

Chitashvili, R. (1983). Martingale Ideology in the Theory of Controlled Stochastic Processes. In: Prokhorov, J. and Itô, K. (eds.), *Probability Theory and Mathematical Statistics*, vol. 1021 of *Lecture Notes in Math.*, pp. 73–92. Springer Berlin.

Coculescu, D., Jeanblanc, M. and Nikeghbali, A. (2008). Default Times, non Arbitrage Conditions and Change of Probability Measures. Available at: arXiv:0812.4064v1.

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC.

Crisan, D. and Manolarakis, K. (2010). Solving Backward Stochastic Differential Equations Using the Cubature Method: Application to Nonlinear Pricing. *Progress in Analysis and its Applications*, pp. 389–397. Proceedings of the 7th International Isaac Congress.

Delbaen and Schachermayer, W. (1994). A General Version of the Fundamental Theorem of Asset Pricing. *Math. Annal.*, vol. 300, no. 3, pp. 426–520.

El Karoui, N., Peng, S. and Quenez, M. (1997). Backward Stochastic Differential Equations in Finance. *Math. Finance*, vol. 7, no. 1, pp. 1–71.

El Karoui, N. and Quenez, M.C. (1995). Dynamic Programming and Pricing of Contigent Claims in Incomplete Markets. *SIAM J. Control Optim.*, vol. 33, no. 1, pp. 29–66.

Frei, C. and Reis, G.D. (2011 Feb). A financial market with interacting investors: does an equilibrium exist? *Math. Financ. Econ.*, pp. 1–22. ISSN 1862-9679. 10.1007/s11579-011-0039-0.

Frei, C.M. (2009). *Exponential Utility Indifference Valuation: Correlation, Semimartingales, BSDEs, Convergence*. Ph.D, ETH Zurich.

Hodges, S.D. and Neuberger, A. (1989). Optimal Replication of Contingent Claim under Transaction Costs. *Revue Futures Markets*, vol. 8, pp. 222–239.

Hu, Y., Imkeller, P. and Müller, M. (2005). Utility Maximization in Incomplete Markets. *Ann. Appl. Probab.*, vol. 15, no. 3, pp. 1691–1712.

Imkeller, P. and Reis, G.D. (2010). Path Regularity and Explicit Convergence Rate for BSDE with Truncated Quadratic Growth. *Stochastic Process. Appl.*, vol. 120, no. 3, pp. 348–379.

Imkeller, P., Reis, G.D. and Zhang, J. (2010 April). Results on Numerics for FBSDE with Drivers of Quadratic Growth. In: Alexander Chiarella, C.N. (ed.), *Contemporary Quantitative Finance*, p. 440. Springer. Essays in Honour of Eckhard Platen.

Imkeller, P., Réveillac, A. and Zhang, J. (2011 April). Solvability and Numerical Simulation of BSDEs Related to BSPDEs with Applications to Utility Maximization. To appear in International Journal of Theoretical and Applied Finance DOI No: 10.1142/S0219024911006437.

Jeanblanc, M. and Cam, Y.L. (2009). Progressive Enlargement of Filtrations with Initial Times. *Stochastic Process. Appl.*, vol. 119, no. 8, pp. 2523–2543.

Jeulin, T. (1979). Grossissement d'une Filtration et Applications. *Séminaire de Probabilité XIII, Lecture Notes in Math.*, vol. 721, pp. 574–609.

Jeulin, T. and Yor, M. (1978). Grossissement d'une Filtration et Semimartingales: Formules Explicites. *Séminaire de Probabilité XIII, Lecture Notes in Math.*, vol. 649, pp. 78–97.

Karatzas, I. and Shreve, S.E. (1988). *Brownian Motion and Stochastic Calculus*. 2nd edn. Springer.

Kazamaki, N. (1994). *Continuous Exponential Martingales and BMO*. 1st edn. Volume 1579 of Lecture Notes in Math. Springer-Verlag.

Kobylanski, M. (1997). Résultats d'Existence et d'Unicité pour les Equations Différentielles Stochastiques Rétrogrades avec Générateurs à Croissance Quadratique. *C. R. Math. Acad. Sci. Paris*, vol. 324, no. 1, pp. 81–86.

Kobylanski, M. (2000). Backward Stochastic Differential Equations and Partial Differential Equations with Quadratic Growth. *Ann. Probab.*, vol. 28, no. 2, pp. 558–602.

Korn, R. (2003). The Martingale Optimality Principle: The Best you can do is Enough. *Wilmott*, vol. 1, pp. 61–67.

Korn, R. and Menkens, O. (2005). On worst-case Investment with Applications in Finance and Insurance Mathematics. In: Deuschel, J.-D. and Greven, A. (eds.), *Interacting Stochastic Systems*, pp. 397–407. Springer Link, New York.

Lazrak, A. and Quenez, M.C. (2003). A Generalized Stochastic Differential Utility. *Math. Oper. Res.*, vol. 28, no. 1, pp. 154–180.

Ma, J., Protter, P. and Yong, M. (1994). Solving Forward-Backward Stochastic Differential Equations Explicitly a Four-Step Scheme. *Probab. Theory Related Fields*, vol. 98, no. 3, pp. 339–359.

Makasu, C. (2009). A note on FBSDE characterization of mean exit times. *C. R. Math. Acad. Sci. Paris*, vol. 347, no. 15-16, pp. 965–969. ISSN 1631-073X. Available at: `http://dx.doi.org/10.1016/j.crma.2009.06.006`

Mania, M. and Tevzadze, R. (2008). Backward Stochastic Partial Differential Equations Related to Utility Maximization and Hedging. *J. Math. Sci.*, vol. 153, no. 3, pp. 291–380. ISSN 1072-3374. 10.1007/s10958-008-9129-9.

Mansuy, R. and Yor, M. (2006). *Random Times and Enlargements of Filtrations in a Brownian Setting*. 1st edn. Volume 1873 of Lecture Notes in Math. Springer.

Matchie, L. (2010 Dec). *Cubature Methods and Applications to Option Pricing*. Master's thesis, University of Stellenbosch.

Morlais, M.-A. (2010). A New Existence Result for Quadratic BSDEs with Jumps with Application to the Utility Maximization Problem. *Stochastic Process. Appl.*, vol. 120, no. 10, pp. 1966 – 1995. ISSN 0304-4149.

Musiela, M. and Zariphopoulou, T. (2010). Stochastic Partial Differential Equation and Portfolio Choice. In: *C. Chiarella and A. Novikov*. Springer.

Nualar, D. (1995). *The Malliavin Calculus and Related Topics* . Probability and its Applications, New York, Berlin.

Pardoux, E. (1996). Backward Stochastic Differential Equations and Viscosity Solutions of Systems of Semilinear Parabolic and Elliptic PDEs of Second Order. *Progr. Probab.*, vol. 1, pp. 55–61. Stochastic Analysis and related Topics, VI.

Pardoux, É. and Peng, S.G. (1990). Adapted Solution of a Backward Stochastic Differential Equation. *Systems Control Lett.*, vol. 14, no. 1, pp. 55–61. ISSN 0167-6911.

Pham, H. (2010 May). Stochastic Control under Progressive Enlargement of Filtrations and Applications to Multiple Defaults Risk Management. *Stochastic Process. Appl.*, vol. 120, no. 9, pp. 1795–1820.

Quenez, M.C. (1993). *Méthodes de Contrôle Stochastique en Finance*. Ph.D, Université Pierre et Marie Curie.

Reis, G.D. (2010 May). *On Some Properties of Solutions of Quadratic Growth BSDE and Applications in Finance and Insurance*. Ph.D, Humboldt-Universität zu Berlin.

Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*. Springer.

Richou, A. (2010). *Étude Théorique et Numérique des Équations Différentielles Stochastiques Rétrogrades*. Ph.D, Université de Rennes 1.

Rogers, L.C.G. and Williams, D. (1987). *Diffusion, Markov Processes, and Martingales*. 2nd edn. John Wiley and Sons.

Rong, S. (2007). BSDEs with Jumps and with Quadratic Growth Coefficients and Optimal Consumption. In: Chuong, N.M. (ed.), *Harmonic, Wavelet and p-Adic Analysis*, pp. 343–361. World Scientific Publishing Co.

Runggaldier, W. (2003). Jump-Diffusion Models. In: Ziemba, W. (ed.), *Handbooks in Finance, Book 1*, pp. 169–209. Elesevier/North-Holland.

Schroder, M. and Skiadas, C. (1999). Optimal Consumption and Portfolio Selection with Stochastic Differential Utility. *J. Econom. Theory*, vol. 89, no. 1, pp. 68 – 126. ISSN 0022-0531.

Sung, J. and Wan, X. (2010). Equilibrium Equity Premium, Interest Rate and the Cost of Capital in a Moral-Hazard Economy. Available at SSRN: http://ssrn.com/abstract=1570986.

Tangpi, L. (2010 May). Stochastic Control: with Application to Financial Mathematics. Postgraduate Diploma Essay, African Institute for Mathematical Sciences.

Yang, H. and Zhang, L. (2005). Optimal Investment for Insurer with Jump-Diffusion Risk Process. *Insurance Math. Econom.*, vol. 37, no. 3, pp. 615–634.

Yong, J. and Zhou, X.Y. (1999). *Stochastic Control: Hamiltonian Systems and HJB Equations*. Springer New York.

Zariphopoulou, T. (2001). A Solution Approach to Valuation with Unhedgable Risk. *Finance Stoch.*, vol. 5, no. 1, pp. 61–82.